Positive autoregulation of *lag-1* in response to LIN-12 activation in cell fate decisions during *C. elegans* reproductive system development

Katherine Leisan Luo[1], Ryan S. Underwood[2], Iva Greenwald[3,4,*]

[1] Integrated Program in Cellular, Molecular and Biophysical Studies,

[2] Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA

[3] Department of Biological Sciences, Columbia University, New York, NY 10027, USA

[4] Lead contact

* Correspondence: isg4@columbia.edu

SUMMARY STATEMENT

We show that *lag-1* positively autoregulates during LIN-12-mediated cell fate decisions during cell fate specification via consensus LAG-1/CSL binding sites within a conserved High Occupancy Target sequence.

ABSTRACT

During animal development, ligand binding releases the intracellular domain of LIN-12/Notch by proteolytic cleavage to translocate to the nucleus, where it associates with the DNA-binding protein LAG-1/CSL to activate target gene transcription. We investigated the spatiotemporal regulation of LAG-1/CSL expression in *C. elegans* and observed that an increase in endogenous LAG-1 levels correlates with LIN-12/Notch activation in different cell contexts during reproductive system development. We show that this increase is via transcriptional upregulation by creating a synthetic endogenous operon, and identified an enhancer region that contains multiple LAG-1 binding sites (LBSs) embedded in a more extensively conserved high occupancy target (HOT) region. We show that these LBSs are necessary for upregulation in response to LIN-12/Notch activity, indicating that *lag-1* engages in direct, positive autoregulation. Deletion of the HOT region from endogenous *lag-1* reduced LAG-1 levels and abrogated positive autoregulation, but did not cause hallmark cell fate transformations associated with loss of *lin-12*/*Notch* or *lag-1* activity. Instead, later somatic reproductive system defects suggest that proper transcriptional regulation of *lag-1* confers robustness to somatic reproductive system development.

Key words: autoregulation, Notch, *C. elegans*, *lag-1*, CSL, HOT, enhancer

INTRODUCTION

The intracellular domain of LIN-12/Notch is essentially a membrane-tethered transcriptional activator released by proteolytic cleavage after ligand binding, and is conserved in all animals (Greenwald and Kovall, 2013). LIN-12/Notch activation is initiated when a ligand of the Delta/Serrate/LAG-2 (DSL) family binds to the extracellular domain of LIN-12/Notch and triggers two cleavage events that result in release of the LIN-12/Notch intracellular domain. This domain becomes translocated to the nucleus, where it forms a core nuclear complex with a CSL (Cbf1/Su(H)/LAG-1) family DNA binding protein and a Mastermind family protein (Jeffries et al., 2002; Petcherski and Kimble, 2000; Wilson and Kovall, 2006). This core nuclear complex binds to the DNA through CSL, which recognizes a TGGGAA or YRTGRGAA motif (Bailey and Posakony, 1995; Chen et al., 2020; Christensen et al., 1996; Lai et al., 2000a; Lai et al., 2000b; Lecourtois and Schweisguth, 1995; Nellesen et al., 1999). Genome-wide studies suggest that chromatin state and other transcription factors influence which genes containing consensus CSL binding sites are bona fide Notch targets in a particular tissue (Castel et al., 2013; Chan et al., 2017; Pillidge and Bray, 2019; Skalska et al., 2015). Target genes may also be sensitive to the level of activated Notch, which has been found to affect the dynamics of CSL association with binding sites (Castel et al., 2013; Krejci and Bray, 2007).

Genes that are master regulators of cell specification are often positively or negatively autoregulated, a mechanism that improves stability and control of gene circuits (Becskei and Serrano, 2000). Such positive autoregulation has been described in some Notch-mediated decisions. A positive autoregulatory loop of CSL is required for normal mechanosensation in *D. melanogaster* through a conserved element called ASE*;* the initiation but not the maintenance of this autoregulation is

Notch-mediated (Barolo et al., 2000). Further analysis of ASE revealed a combinatorial effect of various conserved regions to produce the final spatiotemporal pattern of the CSL protein Su(H) (Liu and Posakony, 2014). In *C. elegans,* a study of LIN-12/Notch transgenes suggested that *lin-12* is positively autoregulated through a conserved binding site for LAG-1 and autoregulation is important for a cell fate decision during early gonadogenesis (Wilkinson et al., 1994). A recent report identified multiple LAG-1 ChIP-seq peaks in the *lag-1* gene using whole-worm ChIP-seq of L4 larvae; however, a *lag-1* positive transcriptional feedback loop was shown not to be active in the germline (Chen et al., 2020). This observation suggests that if there is meaningful binding of LAG-1 to sites in the *lag-1* gene, it may be for autoregulation during somatic cell fate decisions.

Here, we show that in *C. elegans*, LAG-1 positively autoregulates in response to LIN-12/Notch activation in three different somatic cell fate decisions important for reproductive system development. We further identify an enhancer region containing multiple binding sites for LAG-1 that is necessary and sufficient to mediate autoregulation in these contexts. The binding sites that mediate autoregulation are embedded in a conserved high occupancy target (HOT) region, a stretch of open chromatin that is promiscuously pulled down in ChIP-Seq experiments, including those performed with LAG-1 (Chen et al., 2020; Wreczycka et al., 2019). Deletion of the HOT region from endogenous *lag-1* using CRISPR/Cas9 reduced overall level of *lag-1* transcription and abrogated positive autoregulation, but did not compromise viability or produce overt cell fate transformations characteristic of loss of LIN-12/Notch or LAG-1 activity. However, we find that adult hermaphrodites have temperature-sensitive defects in egg-laying and vulval eversion. We suggest that the HOT region and positive transcriptional

autoregulation of *lag-1* contribute to the robustness of the reproductive system to environmental perturbations.

RESULTS

The level of endogenously-tagged LAG-1 increases in Vulval Precursor Cells when LIN-12 is active

Six Vulval Precursor Cells (VPCs) initially have the potential to adopt one of three fates. The anchor cell (AC) of the gonad produces the LIN-3/EGF ligand, which induces vulval development. LIN-3/EGF activates a canonical EGFR-Ras-ERK cascade in P6.p, thereby specifying the 1º fate and promoting the expression of ligands that activate LIN-12/Notch in the neighboring VPCs, P5.p and P7.p, specifying them to adopt the 2º fate. The remaining VPCs — P3.p, P4.p, and P8.p — adopt the 3º fate, dividing to produce daughters that fuse to the major hypodermal syncytium. EGF signaling begins during the L2 stage, but the fates of the VPCs are not fixed until the L3 stage (Fig. 1A) (de la Cova et al., 2017; Sternberg, 2005).

This sequential signaling process also engages feedback mechanisms to ensure precise and robust spatial patterning. In 1º VPCs, there are mechanisms to reinforce EGFR-Ras-ERK activity and to counter potential LIN-12 activation (Berset et al., 2005; Shaye and Greenwald, 2002; Stetak et al., 2006; Underwood et al., 2017). In 2º VPCs, activation of LIN-12/Notch leads to expression of direct transcriptional target genes that encode negative regulators of EGFR-Ras-ERK signaling (Berset et al., 2001; Yoo et al., 2004) and a microRNA that targets *vav-1*/Vav, a negative regulator of LIN-12/Notch signaling, so as to positively stimulate *lin-12* activity (Yoo and Greenwald, 2005). Because such LIN-12 target genes are preferentially expressed or upregulated in 2º VPCs compared to other VPCs, we refer to this characteristic herein as a "2º fate pattern."

We generated two alleles of *lag-1* that were endogenously tagged with fluorescent proteins using CRISPR/Cas9 (Methods): *lag-1(ar611[lag-1::gfp])* and *lag-1(ar613[lag-1::mKate2])* (Methods) (Fig. 1B and Fig. S1A). The position of the tag at the carboxy-terminus should capture all known isoforms. Both alleles appear phenotypically wild-type and there is no apparent difference in expression between them. Initially, we performed a qualitative analysis of LAG-1::mKate2, which revealed a dynamic expression pattern in the VPCs during vulval induction (Fig. S1A). In order to achieve a more quantitative analysis, and potentially reveal more subtle dynamics, we quantitated LAG-1::GFP fluorescence, as GFP appeared brighter overall than mKate2. We observed that both LAG-1::GFP (Fig. 1C) and LAG-1::mKate2 (Fig. S1A) accumulated to a higher level in P5.p and P7.p in relation to other VPCs (2^o fate pattern). This accumulation pattern is consistent with LIN-12 activation in the VPCs and suggests a positive feedback mechanism.

To test the possibility of positive feedback, we manipulated LIN-12 activity in the VPCs and observed the effect on LAG-1::GFP accumulation. Our results support the interpretation that LIN-12 activity increases the expression and/or stability of LAG-1::GFP in VPCs. First, we used the transgene *arTi102*[*lin*-31p::*lin-12(intraΔP)*] to ectopically express a constitutively active and stable nuclear form of LIN-12, LIN-12(intraΔP), which is known to activate *lin-12* target gene reporters in all VPCs (Deng and Greenwald, 2016; Underwood et al., 2017). The presence of constitutively active LIN-12 resulted in the loss of the 2^o fate pattern. The level of LAG-1::GFP in non-2^o VPCs (P4.p, P6.p and P8.p) was increased to be comparable to the level achieved normally in the 2^o VPCs (P5.p and P7.p) (Fig. 1B,D). We note that animals carrying this transgene produce a functional vulva, indicating that P6.p adopts its normal 1^o fate, consistent with mechanisms that resist LIN-12/Notch when

EGFR is active (Underwood et al., 2017), and suggesting that the increase in LAG-1::GFP is a consequence of LIN-12 activation as opposed to an overt cell fate transformation.

We then reduced *lin-12* activity using two different approaches: the null allele *lin-12(n941)* to eliminate *lin-12* activity, and genetic ablation of the AC using the proximal gonad-specific null allele *hlh-2(ar614)* (Attner et al., 2019) to prevent the production of LIN-12/Notch ligands by P6.p to activate LIN-12(+) in P5.p and P7.p. Both treatments resulted in loss of 2^{o} fate pattern, and we observed uniformly low LAG-1::GFP accumulation in all VPCs (Figs. 1E-H).

We made two additional observations during this analysis. One observation pertains to the level of *lag-1* expression in VPCs where LIN-12 is not active: in a *lin-12(+)* background, the level of LAG-1::GFP in P6.p, where EGFR-Ras-ERK signaling is active, is similar to the level in P4.p and P8.p, where EGFR-Ras-ERK signaling is not active (de la Cova et al., 2017). This result suggests that mechanisms that counter potential LIN-12 activation in the presumptive 1^{o} VPC do not do so by negatively regulating LAG-1 accumulation. The other observation is that both methods used to eliminate *lin-12* activity caused all VPCs to have a lower level of LAG-1::GFP accumulation than 3^{o} VPCs in a *lin-12(+)* background. This observation suggests that a basal level of *lin-12* activity, inferred from genetic data (Greenwald and Seydoux, 1990; Greenwald et al., 1983), helps set the initial level of *lag-1* expression.

The level of endogenously-tagged LAG-1 increases where LIN-12 is active during the AC/VU decision in gonadogenesis and in the early M lineage

To determine if the increase of LAG-1 levels in VPCs where LIN-12 is active is a special property of VPCs, we examined two other well-studied paradigms in which LIN-12 activity specifies cell fate during reproductive system development: the developing gonad and the developing sex musculature. In both cases, we observed a higher level of LAG-1::GFP in cells where LIN-12 is active, suggesting that the inferred positive feedback mechanism may be general in somatic reproductive system development.

In the central region of the developing somatic gonad, four cells in the L2 stage initially have the potential to be the anchor cell (AC) or a ventral uterine (VU) precursor cell (Fig. 2A). When the somatic primordium forms, the two outer cells, called β cells, are rapidly specified as VUs and continue to express LIN-12; the two inner cells, called α cells, interact with each other via LIN-12/Notch to resolve which will be the AC and which will be another VU. The LIN-12/Notch-mediated interaction between the α cells is called the AC/VU decision. During the AC/VU decision, *lin-12* and the gene encoding its ligand, *lag-2*, are initially expressed in all four cells; *lin-12* transcription is maintained in presumptive VUs via autoregulation, and *lag-2* transcription is lost in the presumptive VUs through degradation of its transcriptional activator, HLH-2 (Karp and Greenwald, 2003; Sallee and Greenwald, 2015; Wilkinson et al., 1994). Ultimately, all three specified VUs exhibit high LIN-12 activity, while the specified AC does not (Fig. 2A).

We marked the α and β cells of the developing somatic gonad primordium with GFP using *arTi22*[*hlh-2(prox)*p::GFP::H2B]; this marker is initially expressed in all four cells with AC potential, and then becomes progressively restricted from

presumptive VUs until it is expressed only in the presumptive AC (Sallee and Greenwald, 2015). GFP expression therefore allowed us to examine the expression of LAG-1::mKate2 protein as a function of the AC/VU decision. Initially, LAG-1::mKate2 is undetectable in the α and β cells (Fig. 2B). Later in development, but still prior to resolution of the AC/VU decision, LAG-1::mKate2 is present at similar levels in the α and β cells (Fig. 2C). After resolution of the AC/VU decision, the level of LAG-1::mKate2 was higher in the three VUs than in the AC, suggesting that LAG-1 is preferentially expressed in cells with high LIN-12 activity (Fig. 2D). A similar expression profile was observed for the LAG-1::GFP allele (Fig. S2).

The postembryonic blast cell M is present at hatching, and undergoes a dorsal-ventral division in the L1 stage; the dorsal lineage is specified through TGF-β signaling, while the ventral lineage is specified by LIN-12/Notch, via a ligand made by adjacent ventral hypodermal cells (Foehr and Liu, 2008). The stages of the lineage are designated based on the number of descendants. LIN-12 protein is present both dorsal and ventral lineages, starting at the “4-M” stage and continuing through the “18-M” stage, but *lin-12* activity is restricted to the ventral lineage because signal transduction is activated by LIN-12/Notch ligands present ventrally but not dorsally (Foehr and Liu, 2008) (Fig. 2E). *lin-12* activity is required to specify the fates of the sex myoblasts (SMs), which arise in the ventral lineage at the 18-M stage, and migrate to the midbody, flanking the gonad, where later they generate egg-laying vulval muscles. Their lineal cognates in the dorsal lineage are coelomocytes. We observed that LAG-1::GFP is preferentially expressed in the ventral descendants of the M-cell from 4-M to 18-M compared to the dorsal homologs (Fig. 2F).

Together with the observations in VPCs, the analysis of LAG-1::GFP expression in the developing gonad and sex muscle paradigms suggest that preferential LAG-1::GFP expression in cells with LIN-12 signaling may be a general feature of *lag-1* regulation in the somatic reproductive system.

An endogenous "knock-in" transcriptional reporter revealed that the correlation of LAG-1 expression with LIN-12 activity is transcriptionally mediated

Approximately 70% of all *C. elegans* mRNAs are trans-spliced by the addition of a 22nt sequence, either “spliced leader 1” (SL1) or “spliced leader 2” (SL2). SL2 is trans-spliced to downstream genes located in naturally occurring operons in the worm genome, with the SL2 acceptor sequence located in the intergenic region between the upstream and downstream genes (Blumenthal, 2012). Synthetic operons can be generated by placing an intergenic region taken from the CEOPX036 operon between two separate coding sequences. This strategy has been used in fosmid-based contexts to make transcriptional reporters by fusing the SL2 acceptor sequence and the coding region for a fluorescent protein 3’ to the gene of interest (Tursun et al., 2009). We adapted that approach to test if the increase in LAG-1::GFP levels observed in cells where LIN-12 is active reflected increased *lag-1* transcription.

We used CRISPR/Cas9 genome editing to construct *lag-1(ar618)*, a bicistronic *lag-1::gfp::sl2acc::nls::tdTomato::nls* allele, where “*sl2acc*” represents the aforementioned intergenic trans-splicing acceptor region from CEOPX036 (see Methods for details). Two separate mRNAs are created via SL2-mediated trans-splicing: one encoding LAG-1::GFP with a heterologous 3’UTR and the other

NLS::tdTomato::NLS with the endogenous *lag-1* 3'UTR (Fig. 3A). If the 2º pattern of LAG-1::GFP we observed in *lag-1(ar611[lag-1::gfp])* reflects regulation at the transcriptional level, then we would expect to observe that both LAG-1::GFP and NLS::tdTomato::NLS exhibit the same 2º pattern in L3 VPCs. However, if regulation occurs post-transcriptionally, we would expect to see differences between the pattern of LAG-1::GFP and NLS::tdTomato::NLS accumulation, as they share transcriptional regulation but differ in their 3'UTR and protein sequences.

We found that in VPCs during the L3 stage, both LAG-1::GFP and NLS::tdTomato::NLS exhibited a 2º pattern (Fig. 3B and 3C), suggesting that the 2º pattern of LAG-1::GFP observed in *lag-1(ar611)* reflects regulation at the level of transcription.

Identification of an enhancer region that recapitulates the expression pattern of endogenous *lag-1*

A simple hypothesis to account for increased transcription of *lag-1* in cells where LIN-12 is active is that there is positive autoregulation of *lag-1* transcription. To test this possibility, we generated a series of transcriptional reporters. The resulting transgenes were generated using CRISPR/Cas9 single-copy insertion techniques at a well-described site on LG I (Methods) to control for positional and copy-number effects between different transgenes. The *gcy-5* minimal promoter [*gcy-5p(min)*] provides a heterologous transcriptional start site, and was chosen because it alone does not drive expression in VPCs but allows expression in VPCs when combined with other VPC gene regulatory elements (Zhang and Greenwald, 2011)(Fig. S4A).

The *lag-1* gene has four predicted isoforms, with the a/b isoforms starting ~9k upstream of isoform d (Fig. 4A). A transcriptional reporter corresponding to the 5'

intergenic region of the a/b isoform did not drive expression in the VPCs (Fig S4B). We therefore analyzed *lag-1* for conserved potential LAG-1 binding sites (LBSs) that would be located in the potential 5' regulatory regions for the c or d isoforms, which also constitute introns for the a/b isoforms. We performed the search using the consensus sequences TGGGAA or YRTGRGAA, which were derived from the original EMSA studies and also match the motif from a recently published LAG-1 ChIP-Seq (Bailey and Posakony, 1995; Chen et al., 2020; Lai et al., 2000a; Lai et al., 2000b; Lecourtois and Schweisguth, 1995; Nellesen et al., 1999). Additionally, these motifs have been shown to predict functional LBSs in *C. elegans* VPCs (Yoo et al., 2004; Yoo and Greenwald, 2005) (see also Methods and Materials). We identified a cluster of nine LBSs within a sequence of 700-800 bp that is conserved among multiple *Caenorhabditis* species (Fig. S3 and Fig. 4A). The high degree of conservation suggested functional importance, and the large number of conserved LBSs made this region especially attractive to analyze as potentially mediating autoregulation.

This LBS-rich conserved region was additionally attractive based on ChIP-seq studies. First, it is embedded within a 2 kb region that was significantly ($p<0.01$) enriched in all available ChIP-Seq tracks in *C. elegans* at the time we began this study, and therefore may be considered a "high occupancy target (HOT) region" (Wreczycka et al., 2019). Second, Chen et al. showed that LAG-1 also binds to this HOT region, strengthening our inference based on predicted, conserved LBSs that it is a potential site of autoregulation (Chen et al., 2020).

We initially analyzed a 2.7 kb region encompassing this conserved region to study the basis for the 2^{o} pattern of *lag-1* transcription (Fig 4A). This enhancer region contains 18 LBSs and extends through to a Hox site (Roiz et al., 2016)

immediately 3' to the HOT region. A transcriptional reporter containing the intact 2.7 kb enhancer combined with *gcy-5(min)* produced a 2º fate pattern (Fig. 4B), faithfully reporting the expression pattern of endogenously-tagged LAG-1 and the "knock-in" artificial operon.

We then performed additional analysis to identify a minimal region that is sufficient to allow for basal expression in all VPCs and upregulation in 2º VPCs. This analysis established the "1.6 kb enhancer" region, which begins at the start of the conserved region and extends to the 3' end of the 2.7 kb region (Fig. 4C). Smaller regions were not sufficient or drove weaker expression (Figs.4G-H). We note that non-conserved LBSs, and sequences other than LBSs both within or outside of the conserved region, may influence *lag-1* expression, but in sum, the evidence indicates that the conserved LBSs in the 1.6 kb enhancer region are the main elements that confer the 2º pattern.

The 1.6 kb enhancer exhibited preferential expression in VUs and in the ventral lineage of the M-cell from 4-M to early 18-M similar to LAG-1::GFP (Figs. 5A and 5C), suggesting it is a candidate for mediating autoregulation in response to LIN-12 activation in these cell contexts as well.

Increased expression of *lag-1* in cells with active LIN-12 depends on LAG-1 binding sites and indicates positive autoregulation

Upregulation of *lag-1* transcription in cells where LIN-12 is active could reflect direct positive autoregulation via LBSs in *lag-1* or be an indirect effect mediated by other transcription factors. To test the requirement for LBSs, we mutated the LBSs from TGGGAA to AGGGAA and YRTGRGAA to YRAGRGAA, mutations shown to disrupt LAG-1/CSL binding in EMSA studies (Christensen et al., 1996) and to disrupt

lin-12 target gene reporter expression *in vivo* in *C. elegans* (Yoo et al., 2004; Yoo and Greenwald, 2005). Importantly, mutating the LBSs in *cis* in these reporters does not disrupt cell fate specification, avoiding the potentially confounding effect of cell-fate transformation that would occur if LIN-12-mediated cell-fate specification is disrupted by depleting *lag-1* in *trans*.

Our mutational analysis suggests that, in this reporter context, the LBSs in the conserved region of the 2.7 kb enhancer appear to be essential for both basal expression and the 2º VPC pattern. The 2.7 kb enhancer contains 18 LBSs, nine of which reside in the conserved region (Fig. 4B). When we mutated the nine LBSs in the conserved region and left the nine LBSs outside of the conserved region intact, in addition to no longer observing a 2º pattern, we were unable to detect any expression from the 2.7kb enhancer in the VPCs during the L3 stage (Fig. 4F). Conversely, when we left the nine LBSs in the conserved region intact and mutated the nine LBSs outside the conserved region, a stereotypical 2º pattern was detected in the L3 stage (Fig. 4E), including basal reporter expression in P4.p, P6.p, and P8.p.

The abrogation of the basal level of expression of *lag-1* in VPCs when the nine LBSs in the conserved region of the 2.7 kb enhancer are mutated was also observed when we prevented LIN-12 signal transduction in VPCs in a *lin-12(0)* mutant or in the absence of an AC. Together, these results suggest that basal level or maintenance of *lag-1* expression is achieved via an autoregulatory mechanism. The source of ligand may be nearby neurons that express *lag-2* (Li and Greenwald, 2010; Takacs-Vellai et al., 2007) or a low level of LIN-3/EGF from the AC that is below the threshold for 1º fate and associated mechanisms that oppose *lin-12* activity (Barkoulas et al., 2013; Berset et al., 2001; Underwood et al., 2017; Yoo et al., 2004).

We next performed mutational analysis using the 1.6 kb enhancer region. When all 13 LBSs were mutated in the context of the 1.6 kb enhancer, the 2º fate pattern was compromised while basal VPC expression was maintained. The level of expression in both P5.p and P7.p was reduced relative to the wild-type enhancer, with the level in P7.p reduced to the level observed in other VPCs (Fig. 4D); the residual elevation in P5.p may reflect other inputs into this enhancer that are normally masked by upregulation via LBSs. Furthermore, this mutant enhancer did not display the characteristic preferential expression in cells where LIN-12 is active in the somatic gonad and M lineages (Figs. 5B, 5D, 5E).

Together, this analysis indicates that upregulation of *lag-1* transcription in cells where LIN-12 is active reflects direct, positive autoregulation.

Evidence for default repression by LAG-1

In Drosophila and mammals, CSL proteins can act as default repressors of Notch target genes when Notch is inactive (reviewed in (Lai, 2002)). Such default repression was not observed in the *C. elegans* germline, based on the absence of overlap between genes displaying elevated expression in the absence of *lag-1* and germline LAG-1 ChIP-seq peaks (Chen et al., 2020), or in VPCs, where loss of *lin-12* activity does not increase LAG-1 expression in VPCs (Fig. 1) and expression from the intact and LBSmut versions of the 1.6 kb enhancer has a similar baseline level in non-2º VPCs (Fig. 4).

However, our analysis suggests default repression in the AC/VU context. Expression of GFP is qualitatively (Figs. 5A and B) and quantitatively (Fig. 5E) greater in the AC when driven by the LBSmut version of the 1.6 kb enhancer than by the intact 1.6 kb enhancer (Fig. 5E). In addition, the level of expression in the AC

and VU appears similar when driven by the LBS mutant enhancer and there is more variability (Fig. 5E), consistent with "noisier" expression and potential dysregulation resulting from loss of LAG-1-modulated transcription. Qualitatively, it appears that default repression may also be observed in the SM mother and CC lineal homologs, but we were unable to quantify expression in these cells due to technical limitations during image processing. Overall, these results are consistent with other findings that CSL is critical for expression of Notch-activated enhancers in certain cells, but removal of CSL also led to a broadening of expression of such enhancers into other cell types (Morel and Schweisguth, 2000).

Deletion of the HOT region encompassing the conserved LBSs abrogates endogenous LAG-1::GFP expression in VPCs and leads to vulval and egg-laying defects

Our reporter analysis suggested that the LBSs within the conserved HOT region were likely to be necessary for expression and positive autoregulation of *lag-1* in the somatic reproductive system. We did not succeed in attempts to use CRISPR/Cas9 to make multiple specific point mutations in this region, so we instead deleted the 2 kb HOT region that encompasses the conserved region of the allele *lag-1(ar611[lag-1::gfp])* (Fig. 6A). Surprisingly, the resulting mutant, *lag-1(ar611ar647[lag-1(ΔHOT)::gfp]),* was homozygous viable. Nevertheless, there were overt phenotypic consequences consistent with defects in development of the reproductive system: *lag-1(ΔHOT)::gfp)* hermaphrodites were egg-laying defective at all temperatures (n = 63/66 at 25º), and displayed a temperature-dependent abnormal vulval eversion (Evl) defect (n = 121/122 at 15º, n= 36/52 at 25º) (Fig. 6C) (Figs. S6A-C). Such defects are also seen in *lin-12* hypomorphs, and may have several different

underlying causes, including incompletely penetrant cell fate transformations (Sundaram and Greenwald, 1993). The defects we see may therefore reflect less effective LIN-12 activity due to reduction of *lag-1* expression, but it remains possible that they reflect other roles of *lag-1* in later reproductive system development.

We looked for evidence of defects in the AC/VU decision and VPC 2º fate specification in *lag-1(ΔHOT)::gfp* hermaphrodites. We used *arIs222[lag-2*p*::*NLS::tagRFP*]* expression to mark the AC, and observed a single AC in the L3 stage of hermaphrodites grown at 25º, indicating that the AC/VU decision had been made normally. LAG-1::GFP fluorescence of *lag-1(ΔHOT)::gfp* appeared to be much dimmer in the somatic gonad when compared with *lag-1::gfp*, but expression in the head at all stages and in L1 rectal cells was not affected (data not shown), consistent with the absence of hallmark Lag phenotypes associated with L1 lethality (Lambie and Kimble, 1991).

To assess 2º VPC fate specification, we first quantitated the fluorescence of LAG-1::GFP in the VPCs of *lag-1(ΔHOT)::gfp* hermaphrodites at 25ºC, as in our analysis above. We found that LAG-1::GFP levels were significantly decreased in all VPCs, below basal levels seen for the intact locus, consistent with observations described above suggesting the basal expression in VPCs is mediated by LIN-12/Notch activation of LBSs located in the HOT region, and that there was no 2º pattern, consistent with loss of positive autoregulation (Figs. 6D and 6E).

The reduction in *lag-1::gfp* expression in *lag-1(ΔHOT)::gfp* hermaphrodites does not compromise the overt specification of the 2º fate. Loss or strong reduction of *lin-12* activity results in a failure of lateral signaling, such that P5.p and/or P7.p adopt the 1º or 3º vulval fate instead of the 2º vulval fate (Greenwald et al., 1983; Sundaram and Greenwald, 1993). Partial failure of lateral signaling can be

assessed by the ectopic expression of 1º fate markers in prospective 2º VPCs even if they otherwise have 2º character (Berset et al., 2001; Yoo et al., 2004). The *arIs222[lag-2*p::NLS::tagRFP*]* marker of AC fate is also a marker of 1º VPC fate, and we did not observe ectopic expression of tagRFP in P5.p and P7.p, suggesting that 2º fate specification occurred normally. A difference in expression between *lag-1(ΔHOT)::gfp* and *lag-1::gfp* was noted later: as the lineage progresses from mid to late L3, ectopic expression of *arIs222[lag-2*p::NLS::tagRFP*]* becomes increasingly evident: 18% of animals at the "Pn.px" stage, and 45% at the “Pn.pxx” stage (Fig. 6F-G). This observation suggests execution of the 2º fate is compromised.

Abnormalities in 2º fate execution suggested by ectopic *lag-2* reporter expression in VPC descendants may contribute to abnormal vulval eversion, but the penetrance of the egg-laying defect in animals grown at 25º is higher than any vulval abnormalities observed. Defects in later diversification of the egg-laying musculature (Hale et al., 2014) as well as defects in terminal features of gonadal and vulval cell types from early or persistent deficit in *lag-1* transcription, may underlie the egg-laying defect as well as contributing to abnormal vulval morphology.

DISCUSSION

This study began with the observation that the accumulation of endogenous LAG-1::GFP becomes higher in cells where LIN-12/Notch is active than in cognate cells with low LIN-12/Notch activity in three important cell fate decisions in reproductive system development. We constructed a synthetic operon, in which the spliced leader sequence *sl2* was used to drive tdTomato expression in tandem with endogenous *lag-1* transcription, and demonstrated that this LAG-1::GFP upregulation was at the level of transcription. We then identified an enhancer rich in

LAG-1 Binding Site (LBS) sequences and, using transgenes, showed that these LBS sequences are required for upregulation in response to LIN-12/Notch activation. We further observed that this enhancer region is embedded in a High Occupancy Target (HOT) region identified in ChIP-seq experiments for many different transcription factors, including LAG-1 (Chen et al., 2020; Wreczycka et al., 2019) .

When we deleted the entire HOT region from the endogenous *lag-1* gene, we saw greatly reduced expression from the locus, as well as failure to upregulate activity in cells where LIN-12 is active. Despite these effects on gene expression, we did not observe hallmark phenotypes associated with loss of *Notch* activity in *C. elegans*: animals were viable, indicating that embryonic cell fate decisions mediated redundantly by *lin-12*/Notch and *glp-1*/Notch were normal; fertile, indicating that *glp-1*/Notch activity in the germline was sufficient to prevent premature entry into meiosis; cell fate marker expression indicated that the AC/VU decision and 2^{o} VPC fate specification mediated by *lin-12/*Notch during reproductive system development were also normal.

However, we observed that somatic reproductive system development was affected, and that hermaphrodites had an egg-laying defect and abnormal vulval eversion. These defects are consistent with defects in later development of the reproductive system, which involves coordination and conjunction of gonad, vulva, and sex muscles, each of which may be individually affected by reduction in *lag-1* activity, such that the overall adult phenotype may reflect additive or cumulative effects.

The temperature-sensitivity of these defects suggests that transcriptional regulation of *lag-1* may contribute to the robustness of reproductive system development. Robustness is defined as the ability of an organism to resist

stochastic, environmental, and genetic perturbations to its development (Felix and Wagner, 2008). This evolutionary mechanism is thought to be vital for maintaining developmental integrity while allowing for cryptic variation to give rise to increasingly fit organisms (Kienle and Sommer, 2013). VPC specification in *C. elegans* has been used as a model to study robustness given that divergence from a normal specification pattern is rare not only in response to environmental stimuli such as temperature, starvation, and differing food sources but also tolerance for genetic variation (Barkoulas et al., 2013; Braendle and Felix, 2008; Grimbert and Braendle, 2014; Zauner and Sommer, 2007). Indeed, the robustness of VPC patterning can withstand significant perturbation of the activity of LIN-3/EGF and LIN-12/Notch (Barkoulas et al., 2013). Notably, even when a marked effect on gene expression was observed, the range of phenotypic variation was still buffered (Braendle and Felix, 2008; (Barkoulas et al., 2013). Thus, it is plausible that the regulated level of *lag-1* activity contributes to robustness to environmental effects for different *lin-12*/Notch-mediated events in somatic reproductive system development.

MATERIALS AND METHODS

C. elegans alleles and transgenes

A full list of strains and genotypes used in this study can be found in Table S1.

lag-1(ar611), *lag-1(ar613)* and *lag-1(ar618)* are endogenously-tagged *lag-1* alleles generated by CRISPR/Cas9 for this study (see below). Strains carrying these alleles were viable and fertile, appear overtly wild-type, and had no detectable cell fate transformations involving the AC/VU decision, VPC fate patterning, or SM specification.

lin-12(n941) is a null allele (Greenwald et al., 1983); because homozygotes are sterile, strains containing *lin-12(n941)* were maintained using the rescuing extrachromosomal array *arEx1442*, which expresses *lin-12(+)* and is marked *myo-3*p::mCherry (Sallee et al., 2015), allowing the identification of homozygous segregants that lack the array based on lack of mCherry expression in body wall muscles.

Transgenes generated during the course of this work are described in the section "*lag-1* Enhancer Analysis" below. The following transgenes were used to mark cells of interest or to ascertain cell fate:

VPCs: *jccTi1[lin-31*p::mCherry::H2B*]* (de la Cova et al., 2017) and *arTi253[lin-31*p::mTurquoise::H2B] (generated by miniMOS (Frokjaer-Jensen et al., 2014) for single-copy genomic insertion of pKL44) are expressed in VPCs during the L2 and L3 stages. The transgene *arIs222[lag-2*p::NLS::tagRFP*]* is expressed in 1° VPCs in response to the inductive signal and therefore marks 1° fate .

Proximal somatic gonad: *arTi112[ckb-3*p:: mCherry::H2B*]* labels all somatic gonad cells in the L2 and early L3 stages, until the somatic gonad blast cells divide and dilute the fluroescent histone (Attner et al., 2019; Tenen and Greenwald, 2019). *arIs222[lag-2*p::NLS::tagRFP*]* is expressed in all four cells with AC potential and resolves to the AC by the mid-L3 stage (Sallee et al., 2015). *arTi22[hlh-2(prox)*p::GFP::H2B*]* was used to label the somatic gonad during L2; expression is initially equivalent in α and β cells, gradually becomes restricted to the α cells only during the AC/VU decision and after β specification to VUs, and finally to the AC only following resolution of the AC/VU decision (Sallee et al., 2017).

M-lineage: *jjIs3900[hlh-8*p::NLS::mCherry*]* marks all M-lineage descendants in the L1 and L2 stages (Shen et al., 2017).

CRISPR/Cas9 genome editing

lag-1(ar611[lag-1::gfp]) was generated through an *unc-22* co-CRISPR strategy, using sgRNAs targeting both the *lag-1* locus and *unc-22* (Kim et al., 2014). The allele was modeled off of a *lag-1* fosmid, whose tag contains a 2xTY1-GFP-frt-2xFLAG (Sarov et al., 2012). A silent mutation was incorporated into the homology repair template to mutate the PAM site. A hygromycin resistance cassette, flanked with loxP sites, was inserted 3' to the *lag-1* 3'UTR.

lag-1(ar613[lag-1::mKate2]) was generated using the self-excising cassette strategy (Dickinson et al., 2015), using a sgRNA that targeted the junction between the last exon of *lag-1* and the *lag-1* 3'UTR, alleviating the need to mutate the PAM site.

lag-1(ar618[lag-1::gfp::sl2acc::nls::tdTomato::nls]) was made using a similar strategy to *lag-1(ar613)*. The NLS::tdTomato::NLS sequence contains an N-terminal SV40 NLS and a C-terminal *egl-13* NLS, followed by TEV-3xFLAG.

lag-1(ar611ar647[*lag-1::gfp(ΔHOT)*])was achieved by injecting preformed Cas9-tracrRNA-crRNA snRNP complexes with the homology repair along pRF4 [*rol-6(su1006)*] (Dokshin et al., 2018) into *lag-1(ar611); jccTi1*[*lin-31*p::mCherry::H2B]. We used PCR-based screening to identify the appropriate mutant, followed by sequence confirmation of the deletion. A list of sequences targeted for editing by sgRNA or crRNA is provided in Table S2, and a list of primers used to generate the homology arms is listed in Table S3. For *ttTi4348* CRISPR/Cas9-medited transgenic insertions, please see (Pani and Goldstein, 2018).

Analysis of LAG-1::GFP expression in VPCs: image acquisition, and quantification

All larvae used for image quantification were grown at 25ºC. L3 larvae were taken at 30 hours after a 2 hour egg-lay and mounted in 12.5 µM levamisole on agarose pads. Image stacks were collected from a spinning disk confocal microscope at 0.25 µm intervals. All photomicrographs used for quantification were taken using a VPC marker, *jccTi1*[*lin-31*p*::*mCherry::H2B*::unc-54 3'*UTR] or *arTi253*[*lin-31*p*::*mTurquoise*::*H2B*::unc-54 3'*UTR] (the latter in the case of quantification of *lag-1(ar618[lag-1::gfp::sl2acc::nls::tdTomato::nls]*). It was noted that there no bleed-through of mTurquoise into the GFP channel, but there was bleed-through of GFP into the CFP channel. An image "blank" was taken at each experiment for flatfielding purposes. A dual camera system (Carl Zeiss) was used to acquire GFP/YFP and mCherry simultaneously. For all images, exposure times used for GFP was 500ms, mCherry was 150ms, YFP was 500ms, RFP from the NLS::tdTomato::NLS was 500ms, and mTurquoise was 1000ms.

Images were processed by a pipeline as described in de la Cova et al. (2017) using ImageAnalysis software from the Covert Lab to identify VPCs P4.p through P8.p (de la Cova et al., 2017; Regot et al., 2014). A custom Matlab script was used to sum the maximum five integrated intensity values of each VPC within an animal to obtain the total fluorescence intensity value. To compare the degree of patterning relative to P6.p, the summed value of each VPC was normalized to the summed value of P6.p within the same animal.

Statistical Analysis of LAG-1::GFP expression in VPCs

Because our sample data did not pass the Shapiro-Wilk test of normality, the non-parametric Wilcoxon Signed Rank was used to compare P5.p and P7.p with P6.p, and Mann-Whitney U tests was used to compare VPCs from different sets of data (e.g. P4.p expression from the 1.6 kb enhancer with P4.p expression from the 1.6 kb enhancer (LBSmut)).

Analysis of LAG-1::GFP and transgene expression in the AC/VU decision and M lineage

We analyzed LAG-1::GFP and wild-type or LBS(mut) transgenes formed from 1.6 kb enhancer reporters in the proximal somatic gonad and in the M lineage. To score expression in the somatic gonad, we marked gonadal cells with *arTi112*[*ckb-3*p::mCherry::H2B] and examined hermaphrodites in the late-L2 stage, after the somatic gonad primordium had formed (Kimble and Hirsh, 1979). Characteristic morphology and position (5R and 5L) in the somatic primordium was used to identify the AC or VUs, and then relative intensities was assessed qualitatively.

To score expression in cells of the M lineage, we marked cells with *jjIs3900*[*hlh-8*p::NLS::mCherry], which was kindly provided by Jun (Kelly) Liu of Cornell University (Shen et al., 2017). Morphology was first determined by counting the number of cells expressing NLS::mCherry to determine the stage of the M-lineage descendants. Only 4-M through 18-M (Foehr and Liu, 2008) (prior to anterior migration of the SM) stages were used for scoring. Preferential expression was then qualitatively assessed through the appropriate GFP or YFP channel (depending on whether the LAG-1::GFP or the enhancers were being assessed) on a spinning disk confocal microscope.

lag-1 enhancer analysis

The enhancer analysis was performed by inserting enhancer transgenes into the defined site on linkage group I, *ttTi4348*, using reagents provided by Ariel Pani and Bob Goldstein (Pani and Goldstein, 2018). This method not only allowed for single-copy analysis but also a controlled genomic environment for quantitation.

The germline injection mixture is as described previously (de la Cova et al., 2017). The homology repair templates were generated using pWZ111, which contained the homology arms specific to *ttTi4348* along with a self-excising cassette that was used to identify successful insertions. The homology repair was injected at 10 ng/ul along with the sgRNA plasmid pAP082 (65 ng/ul), and fluorescent coinjection plasmids pCFJ90[*myo-2*p::mCherry] (2.5 ng/ul) and pGH8[*rab-3*p::mCherry] (10 ng/ul). Injected hermaphrodites were incubated for 2 days at 25^{o}C before a hygromycin solution was added to a final concentration of 250 µg/mL. Surviving F1 and F2s were then assayed for insertion of the transgene.

The 2.7 kb enhancer region was selected to encompass the greatest density of LAG-1 Binding Sites (LBSs, defined as TGGGAA or YRTGRGAA, see Results for more information) in the *lag-1* genomic locus. This region is located in the intron between the first and second exons of isoforms A/B, and in the 5' region upstream of isoforms C and D. This region differs from that of a previous, preliminary enhancer analysis performed using extrachromosomal arrays and a Δ*pes-10* minimal promoter (Choi et al., 2013) in that it contains an additional 800-900 bp to the 3’ end, which we found increases the strength of expression and was crucial for the detailed analysis described herein.

The 2.7 kb enhancer contains 18 LBSs. We performed a BLAT alignment (Kent, 2002) against eight other *Caenorhabditis* species: *C. sinica, C. tropicalis, C. nigoni, C. remanei, C. brenneri, C. briggsae, C. angaria,* and *C. japonica*. Of these species, only *C. japonica* and *C. angaria* failed to align with the enhancer region. The regions that had aligned to the 2.7 kb enhancer were also all found to be in the 5' intergenic region of the *lag-1* homolog of the corresponding *Caenorhabditis* species. None of these regions were located in the intron of any *lag-1* isoforms, as with *C. elegans,* but were primarily located in promoter or intergenic regions, with the exception of *C. tropicalis*, whose conserved region was located in the intron of a gene that was adjacent to its *lag-1* homolog. This gene was found to have no *C. elegans* homologs, nor did it align (using BLASTx) to the annotated coding genes of other *Caenorhabditis* species. We defined the conserved region as the overlapping sequence that was common to all six *Caenorhabditis* species that had possessed an alignment using Clustal Omega.

We identified the HOT region by downloading the HOT region track from Wreckzycka et al. (2019), which had defined a HOT region as the 99th percentile in binding (equivalent to $p<0.01$). We found the track overlapped with the conserved region and was located within the larger 2.7 kb enhancer we had determined to contain the desired cluster of eighteen LBSs that were sufficient to reproduce the LIN-12-dependent expression patterns that were analyzed.

The following transgenes were created from the 2.7 kb enhancer and fused to the 200 bp *gcy-5* minimal promoter to drive expression of 2xNLS::YFP using Gibson assembly cloning (Gibson, 2011):

arSi31: This transgene contained the minimal 200 bp *gcy-5* promoter only.

arSi35: The full 2.7 kb enhancer was used in this transgene.

arSi59: The region 5' to the conserved region was removed from the 2.7 kb enhancer to create the 1.6 kb enhancer.

arSi60: This region is the same as *arSi59* except with all 13 LBSs mutated (TGGGAA→AGGGAA and YRTGRGAA→YRAGRGAA), thus generating the 1.6 kb enhancer (LBSmut)

arSi70: This transgene is the same as *arSi59* except all the LBSs not located in the HOT region were mutated (TGGGAA→AGGGAA and YRTGRGAA→YRAGRGAA).

arSi53: The conserved region only was used in this transgene.

arSi55: The conserved region with an additional ~200bp on each 5' and 3' end was used to make this transgene.

arSi74: The full 2.7 kb enhancer was used in this transgene with only LBSs located in the non-conserved region mutated.

arSi81: The full 2.7 kb enhancer was used in this transgene with only LBSs located in the conserved region mutated.

Additionally, *arSi86* is a transgene comprised of a 1.4 kb gene-to-gene intergenic region from *lag-1*'s 5' upstream neighbor *cpi-1* down to the ATG of *lag-1*'s first exon. This transgene is followed by a 2xNLS::YFP, but does not contain the 200 bp *gcy-5* minimal promoter, because the *lag-1* promoter would provide its own transcriptional start site. All transgenes generated by insertion into *ttTi4348* are also listed in Table S4. For a list of primers used to make these transgenes along with those of the *lag-1* CRISPR alleles, please see Table S5.

ACKNOWLEDGEMENTS

We thank Jessica Chan for assistance in generating the GFP-tagged *lag-1(ar611)* allele. We also thank Julie Canman and Jun (Kelly) Liu for markers that facilitated this analysis, Tim Schedl and Jian Chen for communicating the results of their LAG-1 ChIP-seq study prior to publication, Nuria Flames and Miren Maicas Iragarai for stimulating discussions, Claire de la Cova for providing initial guidance and instruction on image quantification, Peter Sims and Julie Canman for additional guidance on image analysis, Michelle Attner for helpful discussion and comments on the manuscript, and Claudia Tenen and Claire de la Cova for additional thoughtful discussions. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). Research reported in this publication was supported by the Institute of General Medicine of the National Institutes of Health under award numbers R01 GM114140, R01 GM115718, and R35 GM131746 (to I.G.), and F31 CA177168 (to R.S.U.), and by the National Cancer Institute under award number F31 CA236516 (to K.L.L.).

Attner, M. A., Keil, W., Benavidez, J. M. and Greenwald, I. (2019). HLH-2/E2A Expression Links Stochastic and Deterministic Elements of a Cell Fate Decision during C. elegans Gonadogenesis. *Current biology : CB* **29**, 3094-3100 e3094.

Bailey, A. M. and Posakony, J. W. (1995). Suppressor of hairless directly activates transcription of enhancer of split complex genes in response to Notch receptor activity. *Genes & development* **9**, 2609-2622.

Barkoulas, M., van Zon, J. S., Milloz, J., van Oudenaarden, A. and Felix, M. A. (2013). Robustness and epistasis in the C. elegans vulval signaling network revealed by pathway dosage modulation. *Dev Cell* **24**, 64-75.

Barolo, S., Walker, R. G., Polyanovsky, A. D., Freschi, G., Keil, T. and Posakony, J. W. (2000). A notch-independent activity of suppressor of hairless is required for normal mechanoreceptor physiology. *Cell* **103**, 957-969.

Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature* **405**, 590-593.

Berset, T., Hoier, E. F., Battu, G., Canevascini, S. and Hajnal, A. (2001). Notch inhibition of RAS signaling through MAP kinase phosphatase LIP-1 during C. elegans vulval development. *Science* **291**, 1055-1058.

Berset, T. A., Hoier, E. F. and Hajnal, A. (2005). The C. elegans homolog of the mammalian tumor suppressor Dep-1/Scc1 inhibits EGFR signaling to regulate binary cell fate decisions. *Genes & development* **19**, 1328-1340.

Blumenthal, T. (2012). Trans-splicing and operons in C. elegans. *WormBook*, 1-11.

Braendle, C. and Felix, M. A. (2008). Plasticity and errors of a robust developmental system in different environments. *Dev Cell* **15**, 714-724.

Castel, D., Mourikis, P., Bartels, S. J., Brinkman, A. B., Tajbakhsh, S. and Stunnenberg, H. G. (2013). Dynamic binding of RBPJ is determined by Notch signaling status. *Genes & development* **27**, 1059-1071.

Chan, S. K. K., Cerda-Moya, G., Stojnic, R., Millen, K., Fischer, B., Fexova, S., Skalska, L., Gomez-Lamarca, M., Pillidge, Z., Russell, S., et al. (2017). Role of co-repressor genomic landscapes in shaping the Notch response. *PLoS genetics* **13**, e1007096.

Chen, J., Mohammad, A., Pazdernik, N., Huang, H., Bowman, B., Tycksen, E. and Schedl, T. (2020). GLP-1 Notch-LAG-1 CSL control of the germline stem cell fate is mediated by transcriptional targets lst-1 and sygl-1. *PLoS genetics* **16**, e1008650.

Choi, V. N., Park, S. K. and Hwang, B. J. (2013). Clustered LAG-1 binding sites in lag-1/CSL are involved in regulating lag-1 expression during lin-12/Notch-dependent cell-fate specification. *BMB Rep* **46**, 219-224.

Christensen, S., Kodoyianni, V., Bosenberg, M., Friedman, L. and Kimble, J. (1996). lag-1, a gene required for lin-12 and glp-1 signaling in Caenorhabditis elegans, is homologous to human CBF1 and Drosophila Su(H). *Development* **122**, 1373-1383.

de la Cova, C., Townley, R., Regot, S. and Greenwald, I. (2017). A Real-Time Biosensor for ERK Activity Reveals Signaling Dynamics during C. elegans Cell Fate Specification. *Dev Cell* **42**, 542-553 e544.

Deng, Y. and Greenwald, I. (2016). Determinants in the LIN-12/Notch Intracellular Domain That Govern Its Activity and Stability During Caenorhabditis elegans Vulval Development. *G3 (Bethesda)*.

Dickinson, D. J., Pani, A. M., Heppert, J. K., Higgins, C. D. and Goldstein, B. (2015). Streamlined Genome Engineering with a Self-Excising Drug Selection Cassette. *Genetics* **200**, 1035-1049.

Dokshin, G. A., Ghanta, K. S., Piscopo, K. M. and Mello, C. C. (2018). Robust Genome Editing with Short Single-Stranded and Long, Partially Single-Stranded DNA Donors in Caenorhabditis elegans. *Genetics* **210**, 781-787.

Felix, M. A. and Wagner, A. (2008). Robustness and evolution: concepts, insights and challenges from a developmental model system. *Heredity (Edinb)* **100**, 132-140.

Foehr, M. L. and Liu, J. (2008). Dorsoventral patterning of the C. elegans postembryonic mesoderm requires both LIN-12/Notch and TGFbeta signaling. *Dev Biol* **313**, 256-266.
Frokjaer-Jensen, C., Davis, M. W., Sarov, M., Taylor, J., Flibotte, S., LaBella, M., Pozniakovsky, A., Moerman, D. G. and Jorgensen, E. M. (2014). Random and targeted transgene insertion in Caenorhabditis elegans using a modified Mos1 transposon. *Nat Methods* **11**, 529-534.
Gibson, D. G. (2011). Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol* **498**, 349-361.
Greenwald, I. (2012). Notch and the awesome power of genetics. *Genetics* **191**, 655-669.
Greenwald, I. and Kovall, R. (2013). Notch signaling: genetics and structure. *WormBook*, 1-28.
Greenwald, I. and Seydoux, G. (1990). Analysis of gain-of-function mutations of the lin-12 gene of Caenorhabditis elegans. *Nature* **346**, 197-199.
Greenwald, I. S., Sternberg, P. W. and Horvitz, H. R. (1983). The lin-12 locus specifies cell fates in Caenorhabditis elegans. *Cell* **34**, 435-444.
Grimbert, S. and Braendle, C. (2014). Cryptic genetic variation uncovers evolution of environmentally sensitive parameters in Caenorhabditis vulval development. *Evol Dev* **16**, 278-291.
Hale, J. J., Amin, N. M., George, C., Via, Z., Shi, H. and Liu, J. (2014). A role of the LIN-12/Notch signaling pathway in diversifying the non-striated egg-laying muscles in C. elegans. *Dev Biol* **389**, 137-148.
Jeffries, S., Robbins, D. J. and Capobianco, A. J. (2002). Characterization of a high-molecular-weight Notch complex in the nucleus of Notch(ic)-transformed RKE cells and in a human T-cell leukemia cell line. *Molecular and cellular biology* **22**, 3927-3941.
Karp, X. and Greenwald, I. (2003). Post-transcriptional regulation of the E/Daughterless ortholog HLH-2, negative feedback, and birth order bias during the AC/VU decision in C. elegans. *Genes & development* **17**, 3100-3111.
Kent, W. J. (2002). BLAT--the BLAST-like alignment tool. *Genome Res* **12**, 656-664.
Kienle, S. and Sommer, R. J. (2013). Cryptic variation in vulva development by cis-regulatory evolution of a HAIRY-binding site. *Nature communications* **4**, 1714.
Kim, H., Ishidate, T., Ghanta, K. S., Seth, M., Conte, D., Jr., Shirayama, M. and Mello, C. C. (2014). A co-CRISPR strategy for efficient genome editing in Caenorhabditis elegans. *Genetics* **197**, 1069-1080.
Kimble, J. and Hirsh, D. (1979). The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans. *Dev Biol* **70**, 396-417.
Krejci, A. and Bray, S. (2007). Notch activation stimulates transient and selective binding of Su(H)/CSL to target enhancers. *Genes & development* **21**, 1322-1327.
Lai, E. C. (2002). Keeping a good pathway down: transcriptional repression of Notch pathway target genes by CSL proteins. *EMBO reports* **3**, 840-845.
Lai, E. C., Bodner, R., Kavaler, J., Freschi, G. and Posakony, J. W. (2000a). Antagonism of notch signaling activity by members of a novel protein family encoded by the bearded and enhancer of split gene complexes. *Development* **127**, 291-306.
Lai, E. C., Bodner, R. and Posakony, J. W. (2000b). The enhancer of split complex of Drosophila includes four Notch-regulated members of the bearded gene family. *Development* **127**, 3441-3455.
Lambie, E. J. and Kimble, J. (1991). Two homologous regulatory genes, lin-12 and glp-1, have overlapping functions. *Development* **112**, 231-240.
Lecourtois, M. and Schweisguth, F. (1995). The neurogenic suppressor of hairless DNA-binding protein mediates the transcriptional activation of the enhancer of split complex genes triggered by Notch signaling. *Genes & development* **9**, 2598-2608.
Li, J. and Greenwald, I. (2010). LIN-14 inhibition of LIN-12 contributes to precision and timing of C. elegans vulval fate patterning. *Current biology : CB* **20**, 1875-1879.

Liu, F. and Posakony, J. W. (2014). An enhancer composed of interlocking submodules controls transcriptional autoregulation of suppressor of hairless. *Dev Cell* **29**, 88-101.
Morel, V. and Schweisguth, F. (2000). Repression by suppressor of hairless and activation by Notch are required to define a single row of single-minded expressing cells in the Drosophila embryo. *Genes & development* **14**, 377-388.
Nellesen, D. T., Lai, E. C. and Posakony, J. W. (1999). Discrete enhancer elements mediate selective responsiveness of enhancer of split complex genes to common transcriptional activators. *Dev Biol* **213**, 33-53.
Pani, A. M. and Goldstein, B. (2018). Direct visualization of a native Wnt in vivo reveals that a long-range Wnt gradient forms by extracellular dispersal. *Elife* **7**.
Petcherski, A. G. and Kimble, J. (2000). LAG-3 is a putative transcriptional activator in the C. elegans Notch pathway. *Nature* **405**, 364-368.
Pillidge, Z. and Bray, S. J. (2019). SWI/SNF chromatin remodeling controls Notch-responsive enhancer accessibility. *EMBO reports* **20**.
Regot, S., Hughey, J. J., Bajar, B. T., Carrasco, S. and Covert, M. W. (2014). High-sensitivity measurements of multiple kinase activities in live single cells. *Cell* **157**, 1724-1734.
Roiz, D., Escobar-Restrepo, J. M., Leu, P. and Hajnal, A. (2016). The C. elegans hox gene lin-39 controls cell cycle progression during vulval development. *Dev Biol* **418**, 124-134.
Sallee, M. D., Aydin, T. and Greenwald, I. (2015). Influences of LIN-12/Notch and POP-1/TCF on the Robustness of Ventral Uterine Cell Fate Specification in Caenorhabditis elegans Gonadogenesis. *G3 (Bethesda)* **5**, 2775-2782.
Sallee, M. D. and Greenwald, I. (2015). Dimerization-driven degradation of C. elegans and human E proteins. *Genes & development* **29**, 1356-1361.
Sallee, M. D., Littleford, H. E. and Greenwald, I. (2017). A bHLH Code for Sexually Dimorphic Form and Function of the C. elegans Somatic Gonad. *Current biology : CB* **27**, 1853-1860 e1855.
Sarov, M., Murray, J. I., Schanze, K., Pozniakovski, A., Niu, W., Angermann, K., Hasse, S., Rupprecht, M., Vinis, E., Tinney, M., et al. (2012). A genome-scale resource for in vivo tag-based protein function exploration in C. elegans. *Cell* **150**, 855-866.
Shaye, D. D. and Greenwald, I. (2002). Endocytosis-mediated downregulation of LIN-12/Notch upon Ras activation in Caenorhabditis elegans. *Nature* **420**, 686-690.
Shen, Q., Shi, H., Tian, C., Ghai, V. and Liu, J. (2017). The C. elegans Spalt-like protein SEM-4 functions through the SoxC transcription factor SEM-2 to promote a proliferative blast cell fate in the postembryonic mesoderm. *Dev Biol* **429**, 335-342.
Skalska, L., Stojnic, R., Li, J., Fischer, B., Cerda-Moya, G., Sakai, H., Tajbakhsh, S., Russell, S., Adryan, B. and Bray, S. J. (2015). Chromatin signatures at Notch-regulated enhancers reveal large-scale changes in H3K56ac upon activation. *The EMBO journal* **34**, 1889-1904.
Sternberg, P. W. (2005). Vulval development. *WormBook*, 1-28.
Stetak, A., Hoier, E. F., Croce, A., Cassata, G., Di Fiore, P. P. and Hajnal, A. (2006). Cell fate-specific regulation of EGF receptor trafficking during Caenorhabditis elegans vulval development. *The EMBO journal* **25**, 2347-2357.
Sundaram, M. and Greenwald, I. (1993). Genetic and phenotypic studies of hypomorphic lin-12 mutants in Caenorhabditis elegans. *Genetics* **135**, 755-763.
Takacs-Vellai, K., Vellai, T., Chen, E. B., Zhang, Y., Guerry, F., Stern, M. J. and Muller, F. (2007). Transcriptional control of Notch signaling by a HOX and a PBX/EXD protein during vulval development in C. elegans. *Dev Biol* **302**, 661-669.
Tenen, C. C. and Greenwald, I. (2019). Cell Non-autonomous Function of daf-18/PTEN in the Somatic Gonad Coordinates Somatic Gonad and Germline Development in C. elegans Dauer Larvae. *Current biology : CB* **29**, 1064-1072 e1068.
Tursun, B., Cochella, L., Carrera, I. and Hobert, O. (2009). A Toolkit and Robust Pipeline for the Generation of Fosmid-Based Reporter Genes in C. elegans. *PloS one* **4**.

Underwood, R. S., Deng, Y. and Greenwald, I. (2017). Integration of EGFR and LIN-12/Notch Signaling by LIN-1/Elk1, the Cdk8 Kinase Module, and SUR-2/Med23 in Vulval Precursor Cell Fate Patterning in Caenorhabditis elegans. *Genetics* **207**, 1473-1488.

Wilkinson, H. A., Fitzgerald, K. and Greenwald, I. (1994). Reciprocal changes in expression of the receptor lin-12 and its ligand lag-2 prior to commitment in a C. elegans cell fate decision. *Cell* **79**, 1187-1198.

Wilson, J. J. and Kovall, R. A. (2006). Crystal structure of the CSL-Notch-Mastermind ternary complex bound to DNA. *Cell* **124**, 985-996.

Wreczycka, K., Franke, V., Uyar, B., Wurmus, R., Bulut, S., Tursun, B. and Akalin, A. (2019). HOT or not: examining the basis of high-occupancy target regions. *Nucleic acids research* **47**, 5735-5745.

Yoo, A. S., Bais, C. and Greenwald, I. (2004). Crosstalk between the EGFR and LIN-12/Notch pathways in C. elegans vulval development. *Science* **303**, 663-666.

Yoo, A. S. and Greenwald, I. (2005). LIN-12/Notch activation leads to microRNA-mediated down-regulation of Vav in C. elegans. *Science* **310**, 1330-1333.

Zauner, H. and Sommer, R. J. (2007). Evolution of robustness in the signaling network of Pristionchus vulva development. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 10086-10091.

Zhang, X. and Greenwald, I. (2011). Spatial regulation of lag-2 transcription during vulval precursor cell fate patterning in Caenorhabditis elegans. *Genetics* **188**, 847-858.

Figures

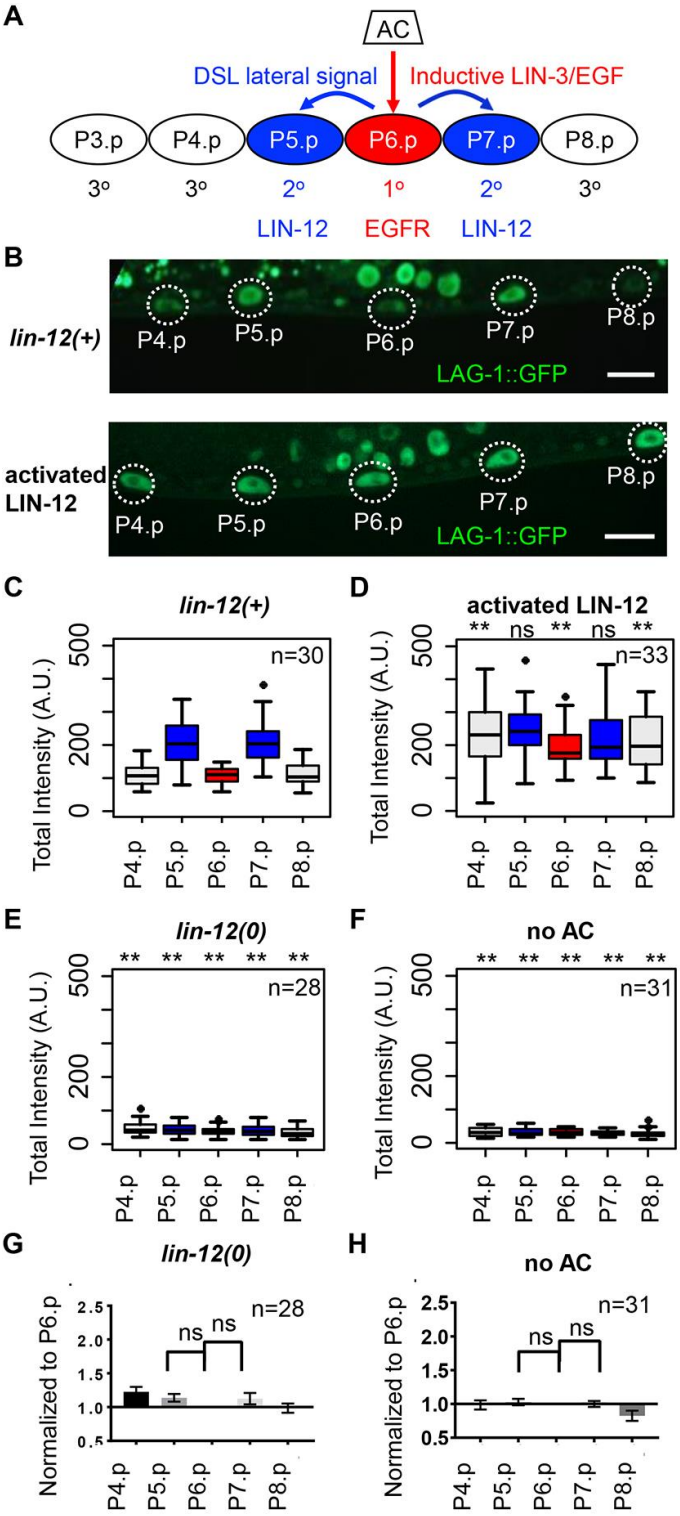

Figure 1. The level of LAG-1::GFP expression in VPCs correlates with *lin-12* activity.

(A) VPC fate patterning (reviewed in (Sternberg, 2005)) is initiated by an EGF-like ligand that activates EGFR in P6.p, the nearest VPC, specifying 1º fate and expression of DSL protein ligands, including LAG-2. The DSL proteins activate LIN-12/Notch in the neighboring VPCs, P5.p and P7.p, leading to target gene expression and 2º fate. Descendants of the 1º and 2º VPCs form the vulva. The other VPCs adopt non-vulval fates: P4.p and P8.p always adopt the 3º fate, which is to produce two daughters that fuse with the major hypodermal syncytium; in about half of wild-type hermaphrodites, P3.p adopts the 3º fate, but alternatively may fuse directly with the syncytium in the L2 stage, so is no longer present as a VPC in the L3 stage. We therefore omit P3.p from our analysis in this study.

(B) Representative fluorescent orthogonal projection of the VPCs of an L3 *lag-1(ar611[lag-1::gfp])* hermaphrodite in a *lin-12(+)* (top) or activated LIN-12 (bottom) background. LAG-1::GFP is visibly brighter in P5.p and P7.p than in other VPCs, the stereotypical 2º fate pattern, in a *lin-12(+)* background and is bright in all VPCs in the presence of activated LIN-12 (quantified in D). Scale bar = 10 μm.

(C) Quantification of LAG-1::GFP fluorescence in VPCs of L3 *lag-1(ar611[lag-1::gfp])* hermaphrodites shows higher levels of LAG-1::GFP in presumptive 2º fate VPCs than in other VPCs, $p<0.001$ by Wilcoxon signed-rank test, which was chosen because of non-normal distribution of the sample populations ($p<0.05$ by Shapiro-Wilk Test of Normality). Here and in all cases below, outliers were determined as values lying beyond 1.5x the interquartile range above the upper quartile or below the lower quartile.

(D-F) All statistical comparisons were performed by Mann-Whitney U Test and compare the level of LAG-1::GFP on a VPC-by-VPC basis in a mutant background

compared to the same VPC in *lin-12(+)* background shown in (C) (** p<0.01, ns not significant).

(D) Quantification of LAG-1::GFP fluorescence with transgenic LIN-12(intraΔP). The presence of constitutively active, stable LIN-12(intraΔP) in VPCs leads to an increase of LAG-1::GFP in non-2º VPCs, equivalent to the level observed in 2º VPCs in the *lin-12(+)* background.

We note that we did not compare the level of LAG-1::GFP in in the intraDP background between different VPCs (e.g. P5.p versus P6.p) because such comparisons are confounded by non-uniform expression of *lin-31*p; values normalized to P6.p are shown in Fig. S1B, and quantification of expression from a *lin-31*p::2xNLS::YFP transgene is shown in and S1C.

(E) Quantification of LAG-1::GFP in a *lin-12(n941)* null background. LAG-1::GFP levels are significantly decreased in all VPCs in the absence of *lin-12* activity compared to *lin-12(+)*

(F) Quantification of LAG-1::GFP in the absence of the Anchor Cell. In *hlh-2(ar614*[*hlh-2(Δprox)*]) hermaphrodites lack an AC, resulting in the failure to induce VPCs to adopt vulval fates. LAG-1::GFP expression is significantly decreased in all VPCs in the absence of the AC compared to *lin-12(+)*

(G) Normalization to P6.p of (E), ns not significant, Wilcoxon signed-rank test.

(H) Normalization of P6.p of (F), ns not significant, Wilcoxon signed-rank test.

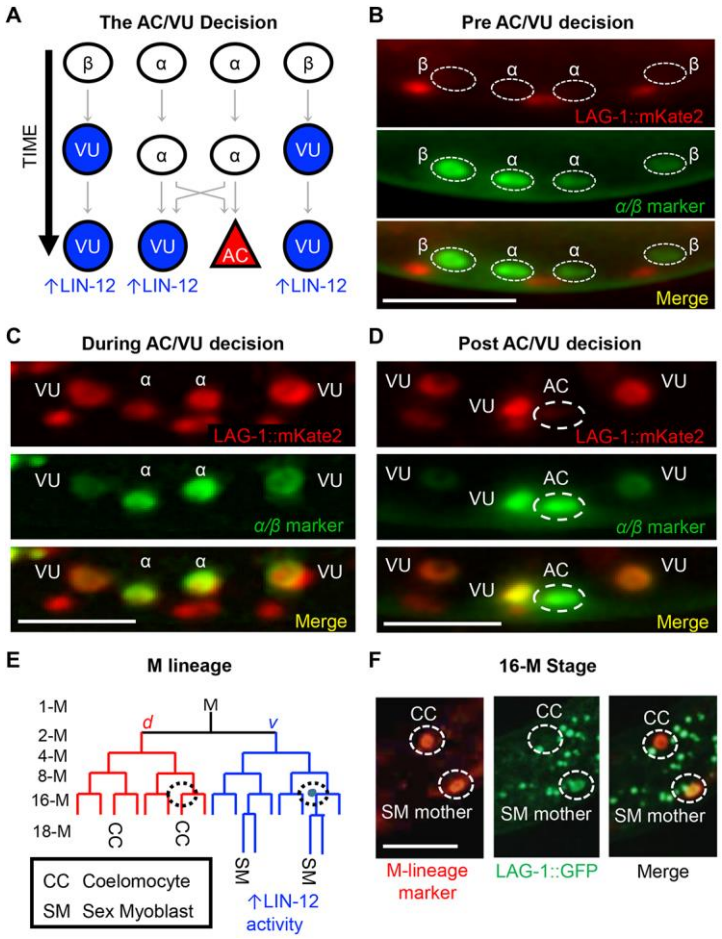

Figure 2. Tagged LAG-1 levels correlate with LIN-12 signaling in lineal cognates during gonad and muscle development.

(A) Each somatic gonad progenitor, Z1 and Z4, gives rise to a pair of α and β cells, all four of which initially have the potential to be an AC (Seydoux et al., 1990). Their developmental potential becomes restricted over time. The β cells commit to the VU fate first; the α cells remain bipotential for longer, and interact with each other to specify one AC and one VU. During the course of their specification, *lin-12* activity increases in the presumptive VUs, and diminishes in the presumptive AC (reviewed in (Greenwald, 2012)). Dynamic influences that impact *lin-12* activity during the AC/VU decision are discussed in (Attner et al., 2019; Sallee et al., 2015).

(B-D) LAG-1::mKate2 from *lag-1(ar613[lag-1::mkate2])* is dynamically patterned during the AC/VU decision in the L2 stage. The α and β cells are marked with GFP from *arTi22[hlh-2(prox)*p::GFP::H2B*]* (Sallee and Greenwald, 2015) and progression is assessed by anatomy. All images shown are max projections.

(B) Early in development, LAG-1::mKate2 is virtually undetectable in the α and β cells (dotted circles).

(C) Later in development, prior to competion of the AC/VU decision, LAG-1::mKate2 is visible in all four α and β cells at similar levels.

(D) Following the AC/VU decision, LAG-1::mKate2 levels remain elevated in the three VU cells while being noticeably reduced in the AC (dashed circle).

(E) LIN-12 is active in ventral M lineage descendants, beginning at the “4M” stage, and is required for sex myoblast (SM) specification (Foehr and Liu, 2008; Greenwald et al., 1983). A coelomocyte (CC) produced in the dorsal branch of the M lineage is the lineal cognate of the parent of the SM in the ventral branch. In the absence of LIN-12, the ventral descendants that normally become SMs instead become CCs. The dashed circles highlight two lineal cognates during 16-M stage, shown in the photomicrographs in (D).

(F) LAG-1::GFP is higher in the SM mother cell compared to its dorsal lineage cognate, the CC, correlating with LIN-12 activity. The M-lineage is labeled by red fluroescence from *jjIs3900*[*hlh-8*p::NLS::mCherry].

Scale bar in all photomicrographs depicts 10 μm.

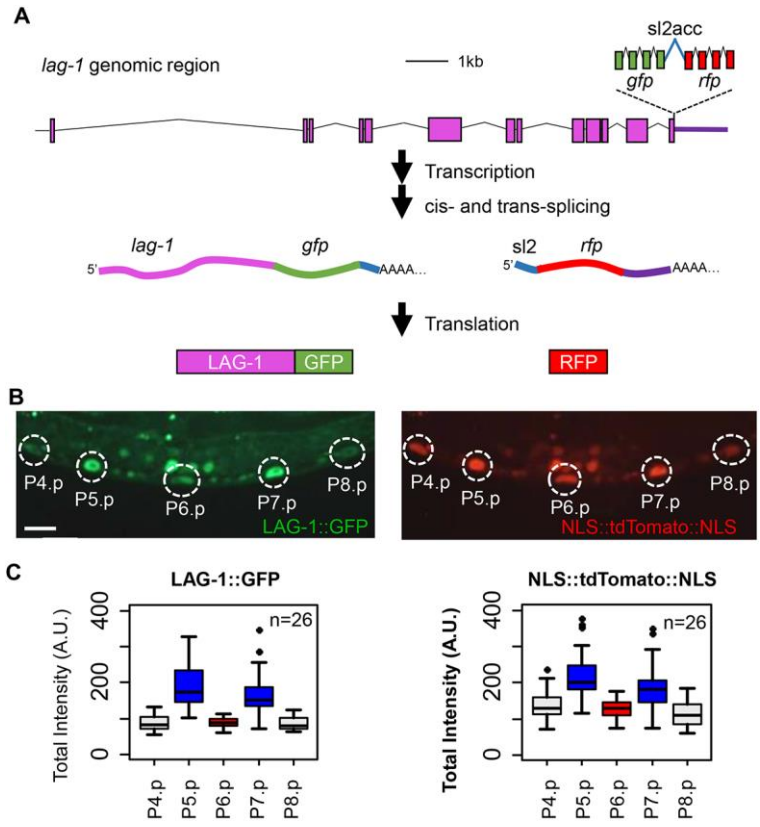

Figure 3. **An endogenous "knock-in" transcriptional reporter reveals transcriptional upregulation of *lag-1* in VPCs in 2º VPCs.**

(A) A schematic of an endogenous *lag-1:gfp::sl2acc::nls::tdTomato::nls* CRISPR/Cas9- "knock-in" that effectively transforms the *lag-1* locus into an operon (Blumenthal, 2012) Two separate mRNA transcripts are produced by SL2 trans-splicing, allowing for independent post-transcriptional regulation of the resultant LAG-1::GFP and NLS::tdTomato::NLS proteins (called RFP in the schematic). If the increase in LAG-1::GFP in 2º VPCs reflects transcriptional upregulation, then LAG-1::GFP and NLS::tdTomato::NLS would both display the 2º VPC fate pattern, with higher levels in P5.p and P7.p than in other VPCs; if there is post-transcriptional regulation of LAG-1::GFP, then a uniform level of NLS::tdTomato::NLS would be seen in all VPCs while LAG-1::GFP would display the 2º VPC fate pattern.

(B) Fluorescent photomicrographs of a *lag-1(ar618[lag-1::gfp::sl2acc::nls::tdTomato::nls])* L3 hermaphrodite. Expression of both LAG-1::GFP (left) and NLS::tdTomato::NLS (right) are observed in the stereotypical 2^o fate pattern.

(C) Quantification of LAG-1::GFP (left) and NLS::tdTomato::NLS (right) fluorescence. For both proteins, the level of fluorescence in P5.p and P7.p is significantly greater than in P4.p, P6.p, and P8.p. All statistics were performed using Wilcoxon signed-rank test using a threshold of $p<0.01$. Scale bar denotes 10 μm.

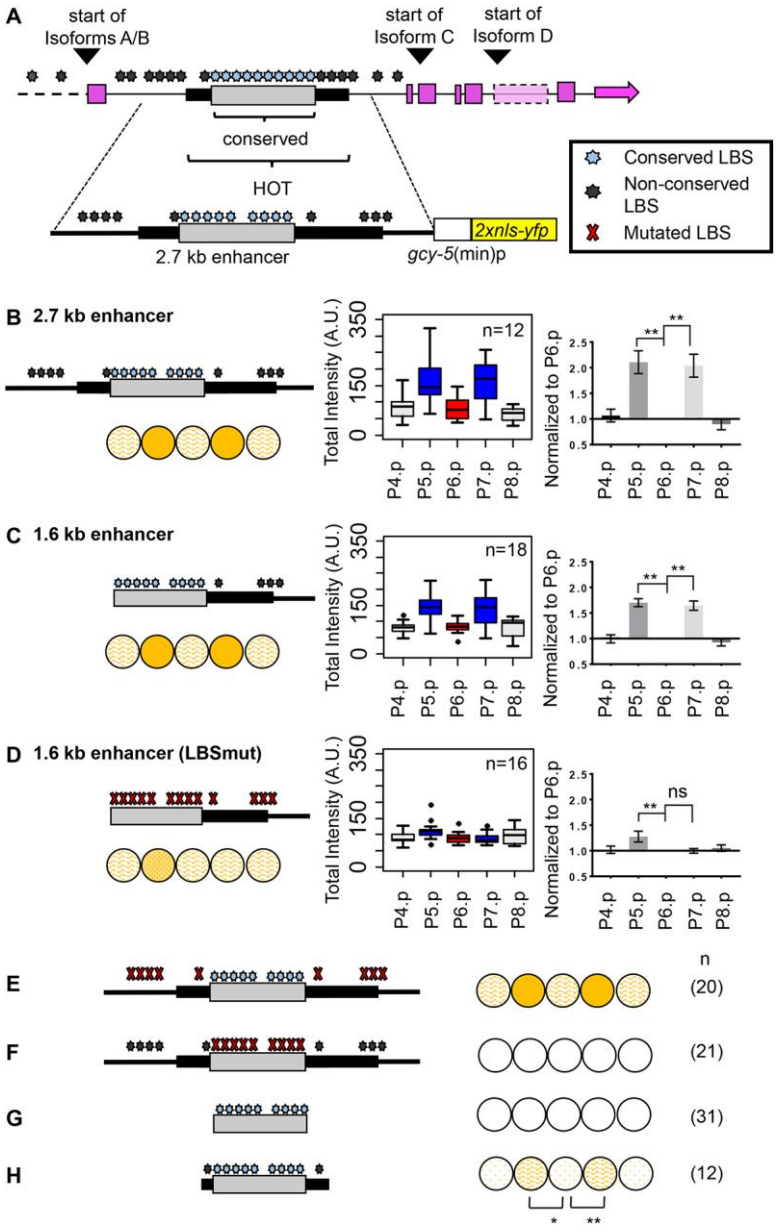

Figure 4. Identification of an enhancer region that recapitulates the expression pattern of endogenous *lag-1*.

(A) Schematic of *lag-1* locus. The start codons of the different isoforms are indicated; the A and B isoforms have the same 5' ATG start codon, and differ by only a single codon at the end of the first exon; the C isoform starts with an exon that is internal to the A and B isoforms; and the D isoform begins with a unique 5' exon that is not shared with any other isoforms (light pink) but shares all downstream exons (not all

shown). A segment represented by the gray bar is conserved among several *Caenorhabditis* species and is centered within a high occupancy target (HOT) region ($p<0.01$) (Wreczycka et al., 2019), represented by the black bar (see Fig. S3). Each potential LAG-1 Binding Site (LBS) is represented by a grey asterisk with LBSs located in the conserved region represented by a blue asterisk. A mutated LBS is denoted by a red-filled X.

(B-H) Transcriptional reporters containing a *lag-1* enhancer region driving *2xnls::yfp*. All reporters were integrated as single-copy insertions into the same site in the genome using CRISPR/Cas9 (see Materials and Methods). B and C show regions that are sufficient to reproduce the 2º VPC fate pattern in the L3 stage.

(B) Quantification of fluorescence from the 2.7 kb enhancer reporter in VPCs in the L3 stage. The 2.7 kb enhancer is sufficient to produce a 2º VPC expression pattern. (Middle) A box plot of the total intensity. (Right) A bar graph with intensity values normalized to P6.p emphasizes the 2º VPC fate pattern as intensities in P5.p and P7.p are both significantly greater than P6.p.

(C) Quantification of fluorescence from the 1.6 kb enhancer region. (Middle) A box plot of the total intensity. (Right) A bar graph with intensity values normalized to P6.p. The total intensity values were not significantly different from total intensity values of corresponding VPCs seen with the 2.7 kb enhancer reporter (Wilcoxon signed-rank test ($p<0.05$).

(D) Mutating all 13 LBS's in the 1.6 kb enhancer context reduces positive autoregulation in VPCs. (Middle) A box plot of the total intensity. (Right) A bar graph with intensity values normalized to P6.p (with statistics performed as above using the Wilcoxon signed-rank test). Upregulation is no longer seen in P7.p; upregulation in P5.p is significantly reduced relative to intact 1.6 kb enhancer ($p<0.01$, Mann

Whitney U Test) but is still observed compared to P6.p. The total intensity of P4.p, P6.p, and P8.p was not significantly different from the corresponding VPCs of the wild-type 1.6 kb enhancer reporter in (C) ($p>0.05$, Mann Whitney U Test), indicating that only upregulation in LIN-12/Notch-specified VPCs was affected.

(E) The 2.7 kb enhancer with LBSs located in non-conserved regions mutated (schematic shown at left) also displays a 2^o VPC fate pattern (right).

(F) The 2.7 kb enhancer with only LBSs located in the conserved region mutated is no longer expressed in VPCs (right).

(G) The conserved region alone is not expressed in VPCs.

(H) Extending the conserved region by 200-300 bp 5' and 3' produced a 2^o VPC pattern, *$p<0.05$, **$p<0.01$, Wilcoxon signed-rank test.

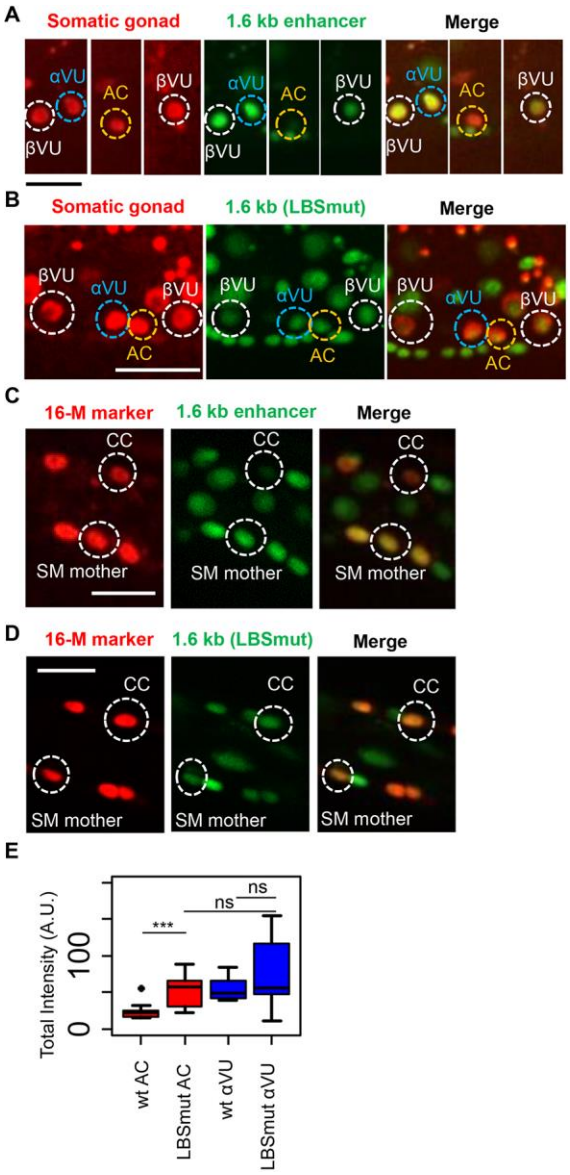

Figure 5. The 1.6kb enhancer shows LBS-dependent preferential expression in cells with high LIN-12 activity compared to lineal cognates in the developing gonad and musculature.

(A) The 1.6kb enhancer is preferentially expressed in the VUs compared to the AC in the somatic gonad (16/16 animals). A representative animal is shown. Each set of panels represents an orthogonal projection, as 2 VUs, the AC, and the third VU are in different planes. The left image displays the somatic gonad marker *arTi112*[*ckb-3*p::mCherry::H2B], while the right image exhibits the enhancer alone.

(B) The 1.6kb (LBSmut) enhancer is not preferentially expressed in the VUs in the somatic gonad post-AC/VU decision, displayed in this representative image (20/20 animals).

(C) The 1.6 kb enhancer is preferentially expressed in the SM mother during the 16-M stage while the corresponding coelomocyte exhibits a lower expression of the enhancer (22/22 animals). Left image shows a merged between the *jjIs3900*[*hlh-8*p::NLS::mCherry] and the 1.6kb enhancer. Right image shows the 1.6 kb enhancer alone.

(D) The 1.6 kb (LBSmut) enhancer is not preferentially expressed in the SM mother compared to the coelomocyte during the 16-M stage (19/19 animals).

(E) Quantification of the AC and the (αVU) for the wild-type and LBSmut forms of the 1.6 kb enhancer. The αVU was chosen for comparison to the AC because of their related lineage histories and more similar positions in the somatic gonad primordium to render the analysis more resistant to imaging distortions arising from capturing cells at different depths. We note that the LBSmut transgene leads to higher LAG-1 level in the AC compared to the wild-type transgene and that there is greater variance in both the AC and αVU for the LBSmut transgene, potentially indicating greater dysregulation of expression with loss of LBSs. ***$p<0.001$, ns not significant by Wilcoxon Signed Rank Test.

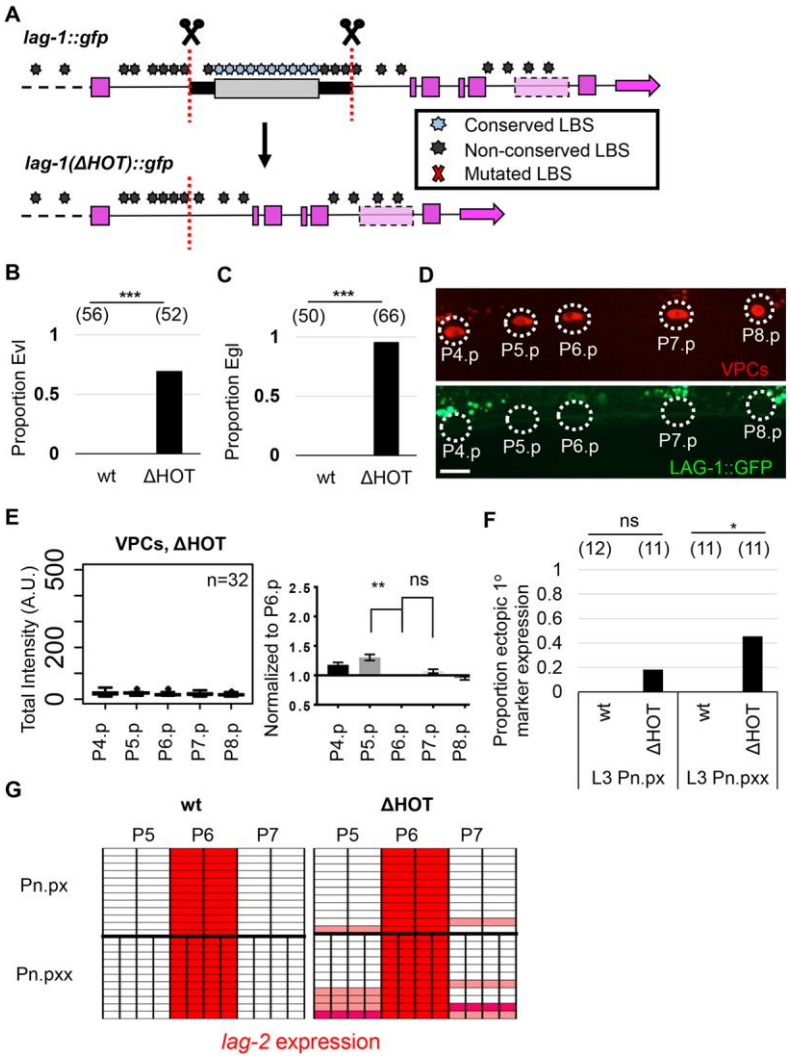

Figure 6. Deletion of the HOT region encompassing conserved LBSs results in egg-laying defective (Egl) and abnormal vulval eversion (Evl) phenotypes.

(A) Deleting the HOT region from the *lag-1(ar611[lag-1::gfp])* resulted in *lag-1(ar611ar647[lag-1(ΔHOT)::gfp]).* In panels C-F, all comparisons are between these two alleles.

(B) Quantification of the Evl phenotype at 25ºC, ***p<0.001, Fisher's Exact Test.

(C) Quantification of the Egl phenotype at 25ºC, ***p<0.001, Fisher's Exact Test.

(D) LAG-1::GFP expression is reduced in all VPCs in *lag-1(ΔHOT)::gfp* L3 hermaphrodites. VPCs are marked with *jccTi1*[*lin-31*p::mCherry::H2B] (top) and

LAG-1::GFP expression in the same cells is difficult to visualize; compare with Fig. 1B. Scale bar indicates 10 μm.

(E) Quantitation of LAG-1::GFP in L3 VPCs shows that expression is unpatterned and lower than the baseline observed in *lag-1::gfp* VPCs. Right is normalization to P6.p. ** indicates $p<0.01$ by Wilcoxon signed-rank test.

(F) Ectopic expression of the VPC 1° fate marker *arIs222*[*lag-2*p::2xNLS::tagRFP] in *lag-1(ΔHOT)::gfp* compared to *lag-1::gfp* in descendants of P5.p and P7.p during the L3 stage, *$p<0.05$, ns = not significant, Fisher's Exact Test. Although visible, this ectopic *lag-2* expression was dimmer than *lag-2* expression in cells descended from the 1° VPC, P6.p.

(G) Another representation of (F) to show distribution and relative brightness of ectopic *lag-2* expression in VPCs.

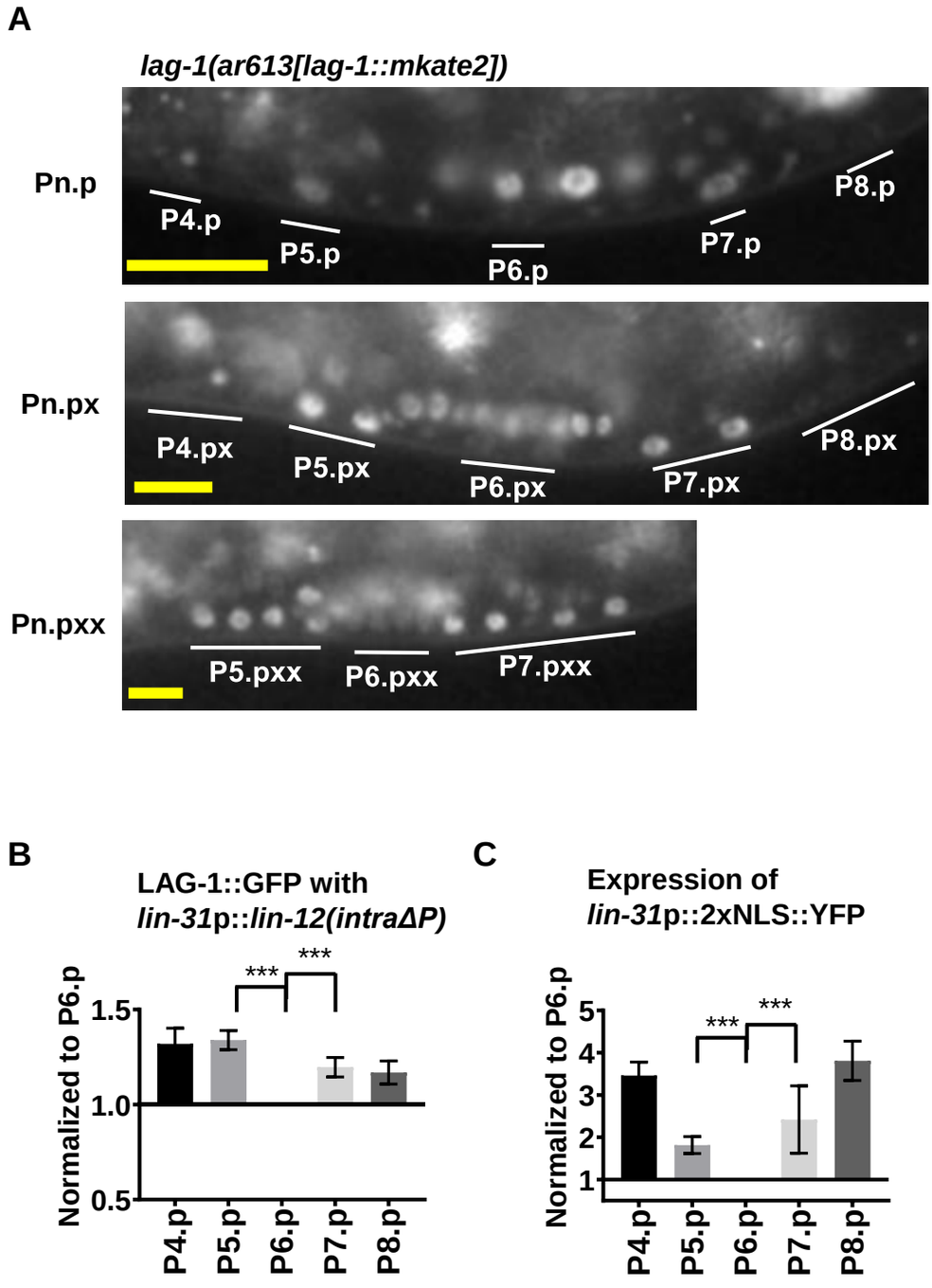

Figure S1. *lag-1(ar613[lag-1::mKate2])* expression, and additional data relevant to Figure 1.

(A) Representative image of *lag-1(ar613[lag-1::mKate2])* from L3 Pn.p stage to L3 Pn.pxx stage. VPCs are underlined and labeled. Scale bar (yellow) = 10 μm.

(B) LAG-1::GFP expression in the presence of *arTi102[lin-31*p::*lin-12(intraΔP)]* (data from Fig. 1D), here normalized to P6.p. Although expression is significantly higher in P5.p and P7.p compared to P6.p, this result is difficult to interpret because of patterned variation in expression that is evident upon quantification of a *lin-31*p transgene in (C). ***p<0.01

(C) Quantification of expression from *arTi88[lin-31*p::2xNLS::YFP*]* during the L3 stage shows patterned variation in expression. Expression is significantly higher in P5.p and P7.p compared to P6.p, suggesting that *lin-31*p itself is affected by spatial patterning signals. ***p<0.01

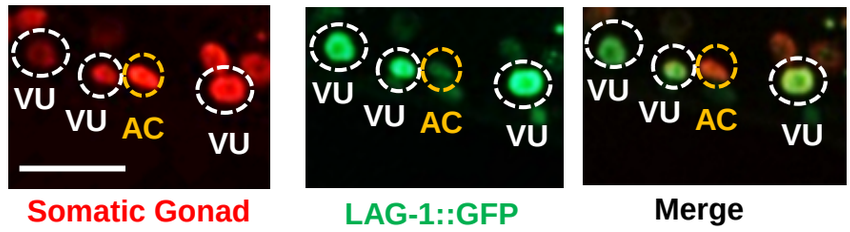

Figure S2. LAG-1::GFP in the somatic gonad
Orthogonal projection of the somatic gonad, in which the AC and VUS have already been specified. (Left) Somatic gonad is marked by *arTi112*[*ckb-3*p::mCherry::H2B]. (Middle) LAG-1::GFP is preferentially expressed in specified VUs. (Right) Merge. Scale bar = 10 µm.

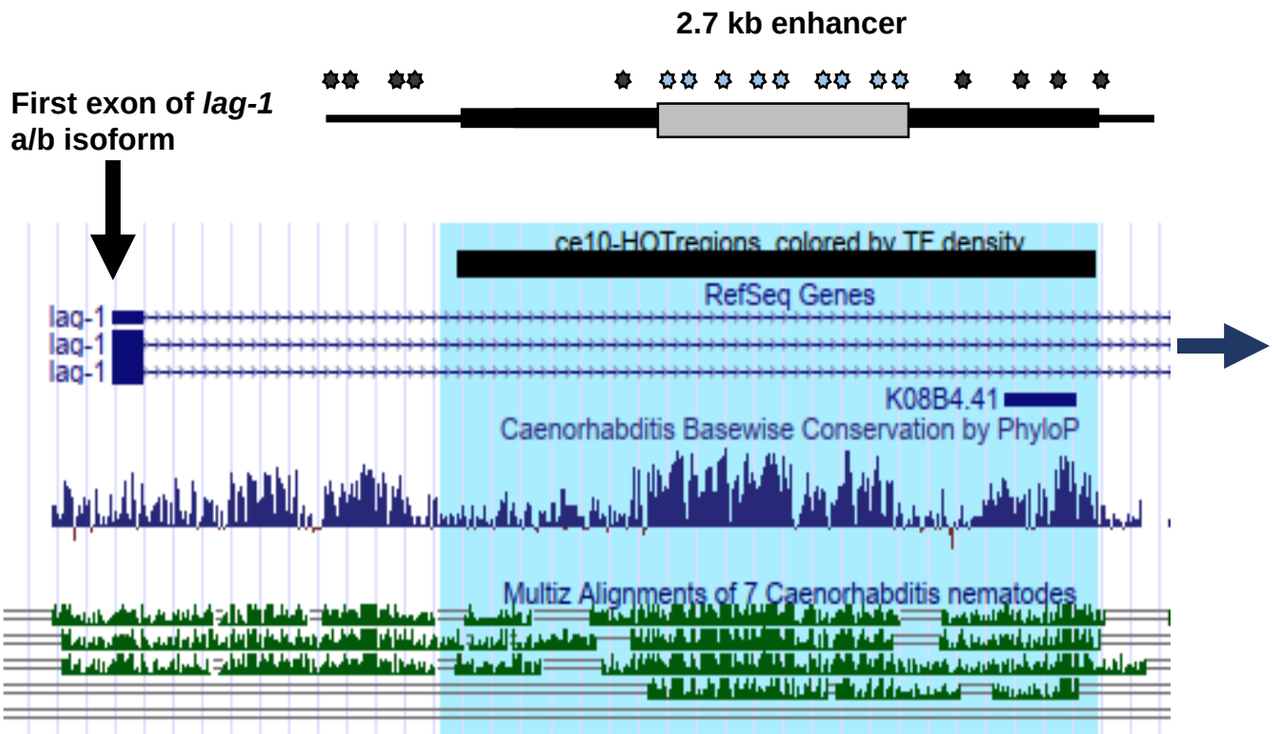

Figure S3. UCSC genome browser view of *lag-1* region used to generate the 2.7 kb enhancer. HOT region track (in black) is provided by Wreczycka, et al. (2019). Conservation tracks are shown in green.

Figure S4

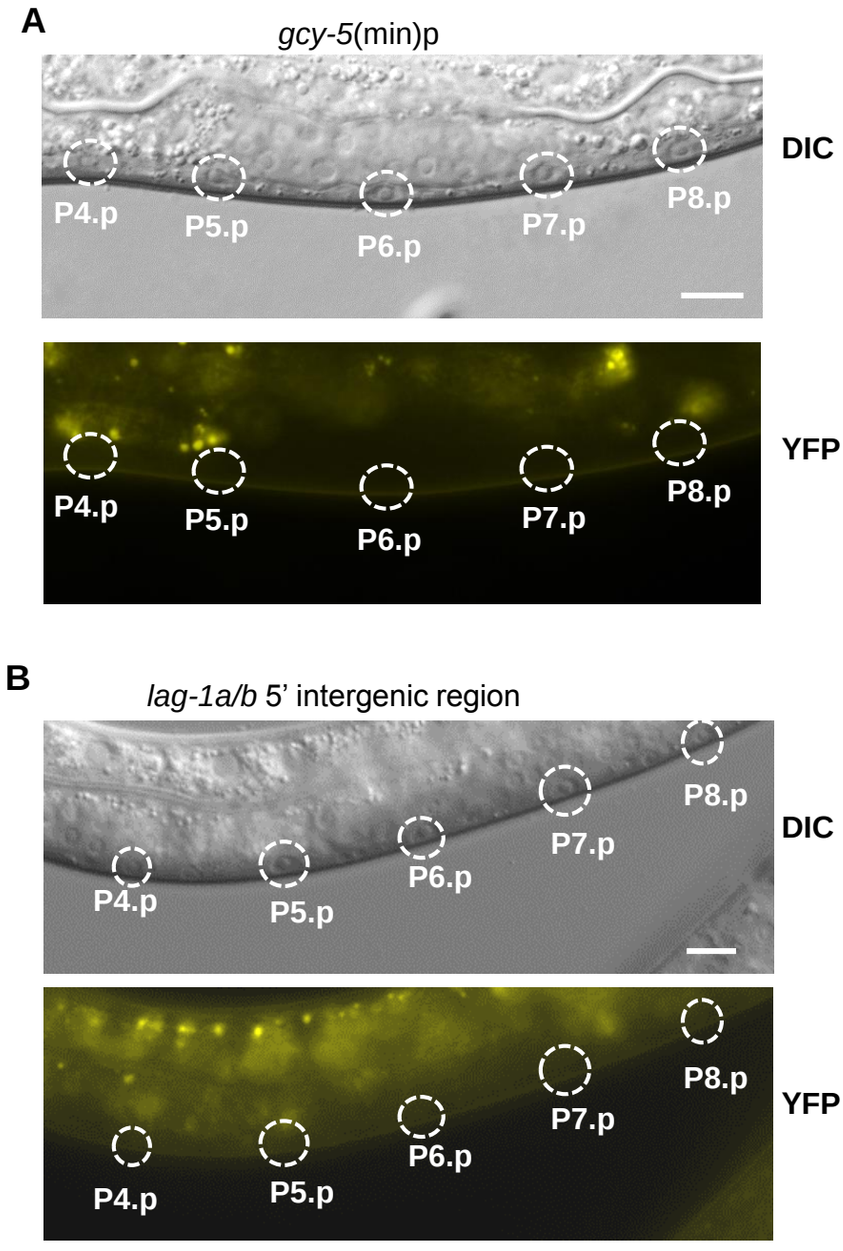

Figure S4. The *gcy-5* minimal promoter and the *lag-1* a/b 5' intergenic region do not result in detectable YFP expression in VPCs.

(A) Representative image of a transgene inserted in the LG I containing only [*gcy-5*(min)p::2xNLS::YFP]. No expression was observed in the L3 VPCs.

(B) Representative image of the intergenic region 5' to the *lag-1* ATG (of isoforms a/b) fused to *2xnls-yfp*. No expression was observed in the L3 VPCs.

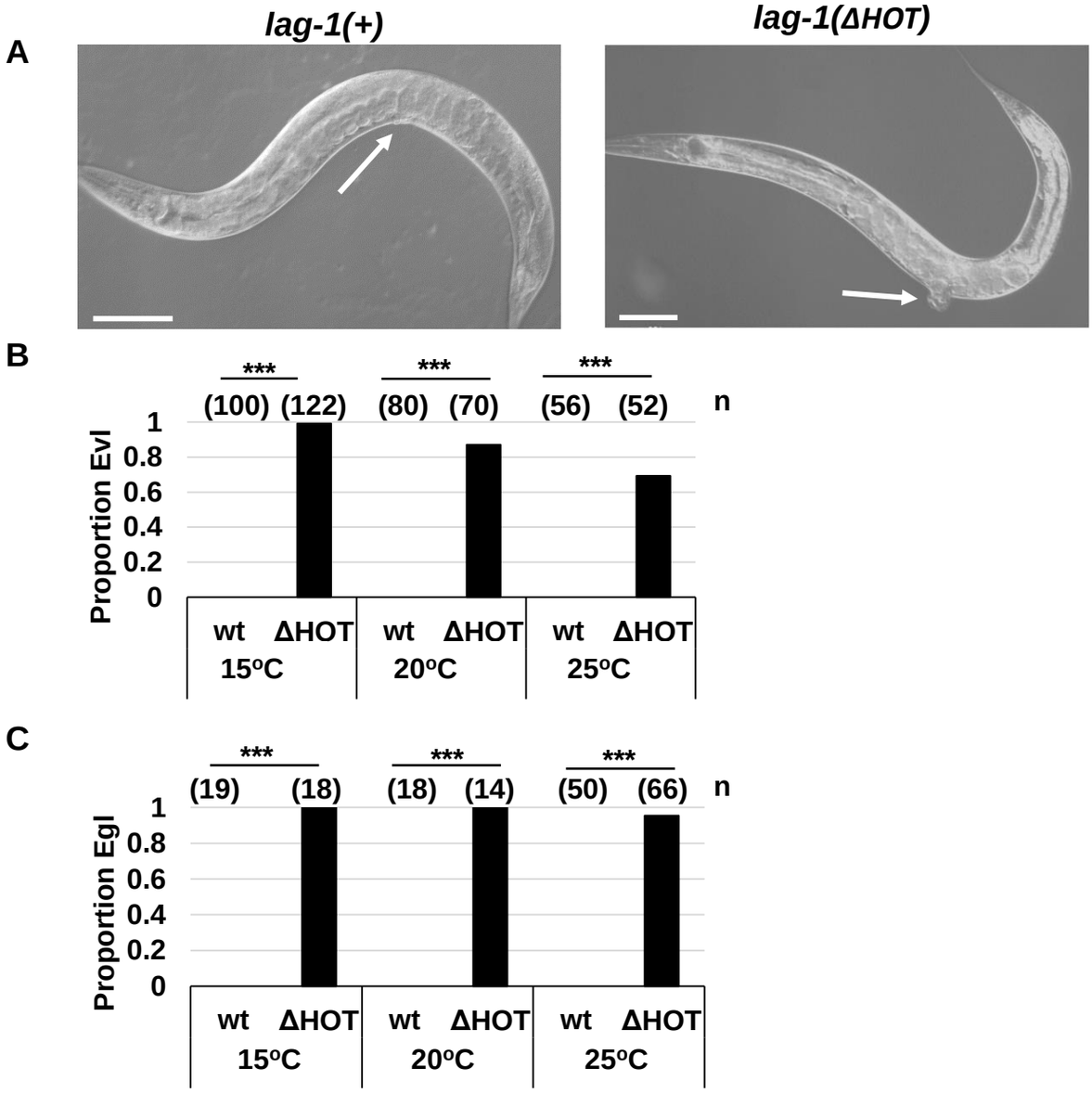

Figure S5. Deletion of the HOT region results in an egg-laying defective (Egl) phenotype and a cold-sensitive abnormal vulval eversion (Evl) phenotype. In addition, no difference in AC marker expression was observed between *lag-1::gfp* and *lag-1(ΔHOT)::gfp*.

(A) Representative images of adult hermaphrodite of *lag-1::gfp* compared to *lag-1(ΔHOT)::gfp* . Scale bar denotes 100 µm.

(B) Temperature dependence of the Evl phenotype. All comparisons between *lag-1::gfp* and *lag-1(ΔHOT)::gfp* were significant ($p<0.001$, Fisher's Exact Test).

(C) Temperature dependence of the Egl phenotype. All comparisons between *lag-1::gfp* and *lag-1(ΔHOT)::gfp* were significant ($p<0.001$, Fisher's Exact Test).

Table S1

Strain List	
Strain	**Genotype**
GS8732	*lag-1(ar611[lag-1::gfp]); jccTi1*[*lin-31*p::mCherry::H2B]
GS8762	*arTi102*[*lin-31*p::*lin-12(intraΔP)*] *lag-1(ar611[lag-1::gfp]); jccTi1*[*lin-31*p::mCherry::H2B]
GS9052	*lin-12(n941); lag-1(ar611[lag-1::gfp]); jccTi1*[*lin-31*p::mCherry::H2B]*; arEx1442*
GS9122	*hlh-2(ar614); lag-1(ar611[lag-1::gfp]); jccTi1*[*lin-31*p::mCherry::H2B]
GS8555	*lag-1(ar613[lag-1::mKate2])*
GS9295	*lag-1(ar611[lag-1::gfp]) jjIs3900*[*hlh-8*p::NLS::mCherry]
GS9047	*lag-1(ar611[lag-1::gfp]); arTi112*[*ckb-3*p::mCherry::H2B]
GS9233	*arSi35; jccTi1*[*lin-31*p::mCherry::H2B]
GS9236	*arSi59; jccTi1*[*lin-31*p::mCherry::H2B]
GS9252	*arSi60; jccTi1*[*lin-31*p::mCherry::H2B]
GS9235	*arSi55; jccTi1*[*lin-31*p::mCherry::H2B]
GS9179	*arSi53*
GS9181	*arSi55*
GS9264	*arSi74*
GS9296	*arSi81*
GS9354	*arSi59; arTi145*[*ckb-3*p::mCherry::H2B]
GS9355	*arSi60; arTi145*[*ckb-3*p::mCherry::H2B]
GS9293	*arSi59; jjIs3900*[*hlh-8*p::NLS::mCherry]
GS9294	*arSi60; jjIs3900*[*hlh-8*p::NLS::mCherry]
GS9397	*lag-1(ar611ar647); jccTi1*[lin-31p::mCherry::H2B]
GS9441	*lag-1(ar611ar647[lag-1(ΔHOT)::gfp]*); *arIs222*[*lag-2*p::2xNLS::tagRFP]
GS9442	lag-1(ar611*[lag-1::gfp]*); *arIs222*[*lag-2*p::2xNLS::tagRFP]
GS8741	*lag-1(ar613[lag-1::mKate2]); arTi22*[*hlh-2(prox)*p::gfp::h2b]

Table S2

Targeting sequence used for CRISPR/Cas9	
Target	**Sequence**
lag-1::gfp insertion	ATGGTGTCGTCTACTCGTC
lag-1 C-terminal insertions	CGAGAGTGGAATCTAGTAAT
lag-1 HOT deletion	TTGAGGTTCCCATGATGCTC, GTATAATCCGTTGAAGATTG

Table S3

Primers for homology template		
Target	**F**	**R**
lag-1::gfp insertion 5' homology arm	TTCAGTTGACGACGAACGTGG	final codon of *lag-1* (*lag-1* Sarov fosmid was used as the template so C-terminal GFP was used for R primer)
lag-1::gfp insertion 3' homology arm	GACATGATGTATCTCGGATTTTGTGGAAC	GCTTGTTGTTCTCATCTCTGCCAC
lag-1 C-terminal insertions 5' homology arm	AGGACAACCGGCGATGTTGAG	GTAATTGGACACAATTCTGCACGGTC
lag-1 C-terminal insertions 3' homology arm	TAGATTCCACTCTCGCGGGATTACTG	GAATCGGATGCGTGGATAGTTGATAATTTATCTG
lag-1 HOT deletion ssODN repair	CGGAAGTACAGCTAAAATGTGTGAGATCTAGGTTAGTTCCCATGATGCTCTGGGCAATTTCGCACATGAC	

Table S4

Transgenes made using *ttTi4348*				
Strain	**Genotype**	***ttTi4348* Transgene**	**Plasmid**	**Figure**
GS9233	***arSi35***; *jccTi1*	lag-1enhancer (2.7kb) + gcy-5(min)p::2xnls-yfp	pKL77	4B
GS8999	***arSi35***			4E
GS9236	***arSi59***; *jccTi1*	lag-1enhancer (1.6kb) + gcy-5(min)p::2xnls-yp	pKL107	4C
GS9293	***arSi59***; *jjIs3900*			5A
GS9354	***arSi59***; *arTi145*			5B
GS9252	***arSi60***; *jccTi1*	lag-1enhancer (1.6kb ΔLBS) + gcy-5(min)p::2xnls-yfp	pKL108	4D
GS9355	***arSi60***; *arTi145*			5C
GS9294	***arSi60***; *jjIs3900*			5D
GS9264	***arSi74***	lag-1enhancer (2.7kb Δnon-conserved LBS) + gcy-5(min)p::2xnls-yfp	pKL131	4F
GS9296	***arSi81***	lag-1enhancer (2.7kb Δconserved LBS) + gcy-5(min)p::2xnls-yfp	pKL110	4G
GS8992	***arSi31***	gcy-5(min)p::2xnls-yfp	pKL78	S4A
GS9353	***arSi86***	lag-1p(gene-to-gene 1.4kb)::2xnls-yfp	pKL139	S4B
GS9235	***arSi55***; *jccTi1*	lag-1enhancer(conserved +400bp) + gcy-5(min)p::2xnls-yfp	pKL103	S4C
GS9250	***arSi70***	lag-1enhancer(1.6kb enhancer Δnon-HOT LBS) + gcy-5(min)p::2xnls-yfp	pKL122	S4D
GS9179	***arSi53***	lag-1enhancer(conserved) + gcy-5(min)p::2xnls=yfp	pKL96	S4E

Table S5

Primers used in enhancer analysis		
Transgene	**F Primer**	**R Primer**
arSi35	TTCCCATCCTAGTTTTTCCCACAC	AAATTCAAATACTGGCATAGAATATATAACTAGTTTTC
arSi59	CGTCATCGTCCTCTGTCGC	AAATTCAAATACTGGCATAGAATATATAACTAGTTTTC
arSi60	CGTCATCGTCCTCTGTCGC	AAATTCAAATACTGGCATAGAATATATAACTAGTTTTC
arSi74	TTCCCATCCTAGTTTTTCCCACAC	AAATTCAAATACTGGCATAGAATATATAACTAGTTTTC
arSi81	TTCCCATCCTAGTTTTTCCCACAC	AAATTCAAATACTGGCATAGAATATATAACTAGTTTTC
arSi31	GCAGATACCAACAAGATTAAAACTTCAAAC	TTTTCATCAGAATAAGTAATTTTTCGAAAACAATAAATAG
arSi86	AACTTTATTTTTAGAAAAGCGAATTTTACCTTCA	CTGAAATTTCTGAATGTTATTTTCATCAATTATAAC
arSi55	AAAATTACATTCCGCACTGCCAG	TTAGGCTTAGTAATGTTGTTTTCTAAGCC
arSi70	CGTCATCGTCCTCTGTCGC	AAATTCAAATACTGGCATAGAATATATAACTAGTTTTC
arSi53	CGTCATCGTCCTCTGTCGC	TTTCTTAGTACTTTTCAATCTTCCCACCAG